Quantitative Economics 2 (2011), 211–249 1759-7331/20110211

Is distance dying at last? Falling home bias in fixed-effects models of patent citations

Rachel Griffith
Institute for Fiscal Studies and University of Manchester

Sokbae Lee
Institute for Fiscal Studies and Seoul National University

John Van Reenen
Centre for Economic Performance, London School of Economics, NBER, and CEPR

We examine the "home bias" of knowledge spillovers (the idea that knowledge spreads more slowly over international boundaries than within them) as measured by the speed of patent citations. We present econometric evidence that the geographical localization of knowledge spillovers has *fallen* over time, as we would expect from the dramatic fall in communication and travel costs. Our proposed estimator controls for correlated fixed effects and censoring in duration models, and we apply it to data on over two million patent citations between 1975 and 1999. Home bias is exaggerated in models that do not control for fixed effects. The fall in home bias over time is weaker for the pharmaceuticals and information/communication technology sectors where agglomeration externalities may remain strong.

Keywords. Fixed effects, home bias, patent citations, knowledge spillovers.

JEL classification. F23, O32, O33.

"When an industry has thus chosen a locality... it is likely to stay there... so great are the advantages... The mysteries of the trade become no mysteries; but are as it were in the air, ... inventions and improvements in machinery, in processes and the general organization of the business have the merits promptly discussed; if one man starts a new idea, it is taken up by others and combined with suggestions of their own..." (Marshall (1890, IV,x,3)).

Rachel Griffith: rgriffith@ifs.org.uk
Sokbae Lee: sokbae@gmail.com
John Van Reenen: j.vanreenen@lse.ac.uk
We would like to thank the European Research Council for research grants and the Economic and Social Research Council for its financial support for the Institute for Fiscal Studies, the Centre for Economic Performance, and the Centre for Microdata Methods and Practice. Sharon Belenzon, Richard Freeman, Jerry Hausman, Bruce Meyer, Peter Thompson, seminar participants at Berkeley, Florida International University, and UCL/IFS, a co-editor, and two anonymous referees have offered useful comments.

DOI: 10.3982/QE59

1. Introduction

The international diffusion of ideas lies at the heart of economic growth and the improvement of the welfare of nations. Unlike most commodities, knowledge is hard to appropriate by its inventors and "spills over" to other agents in the economy. Understanding how knowledge spreads is key to understanding a number of growth enhancing policies (for example, to work out the optimal subsidy to research and development (R&D) or the degree of intellectual property protection). In this paper, we revisit the question of whether geographic proximity plays an important role in the spread of knowledge and, in particular, how this has changed over time. In the popular imagination, the notion of the "death of distance" (Cairncross (1997), Coyle (1997), Friedman (2005)) expresses the idea that information now travels around the globe at rapid speed. Under this view, ideas generated in California spread to Calcutta or Coventry through the Internet, conferences, telephone, and other communication devices at an unprecedented rate, and international boundaries play little role. There is some empirical evidence to support this view (see, inter alia, Keller (2002) and Thompson (2006)).

There are also several counterarguments that suggest that geographical proximity continues to exert a strong influence over knowledge flows. Indeed, in the trade literature, there is little evidence that distance has become any less important for trade flows (e.g., the meta-analysis of Disdier and Head (2008) or Leamer (2007)), and some evidence that its importance may have actually increased (e.g., Evans and Harrigan (2005) and references therein). Distance may still matter if face-to-face interaction is important, even in high-tech sectors, because knowledge is tacit and hard to codify. Globalization may also mean increasing specialization in the technologies where countries have comparative advantage, implying that they have "less to learn" from one another. So ultimately this is an empirical question: Do technology spillovers increase with geographic proximity and has this changed over time?

Figure 1 presents some raw data that is consistent with the view that distance *has* become less important over time for the international transmission of ideas (we discuss the data in much more detail later in the paper). We plot the relative speed of patent citations over time. For example, in the top left panel, consider successful applications to the U.S. Patent Office for inventors living in Germany in an "early" period (1975–1989) on the left and then in a "later" period (1990–1999) on the right. Looking first at the early period, the height of each bar indicates how much slower foreign inventors were in being first to cite German inventors relative to other German inventors. So American inventors were about 40% slower in citing Germans patents than Germans themselves and the French were about 25% slower. The fact that the bars are almost all positive suggests the well known phenomenon of home bias in ideas: Germans are quicker at citing other Germans, Britons are quicker at citing other Britons, and so forth. What is more interesting about Figure 1 is how home bias has *changed over time*. On average, the bars in the later period are lower than the bars in the earlier period, suggesting that home bias in ideas has fallen, consistent with some "death of distance" ideas. In the post-1990 period, Americans are only about 20% slower in citing Germans and the French are only about 10% slower in citing Germans than the Germans themselves. Table 1 holds the

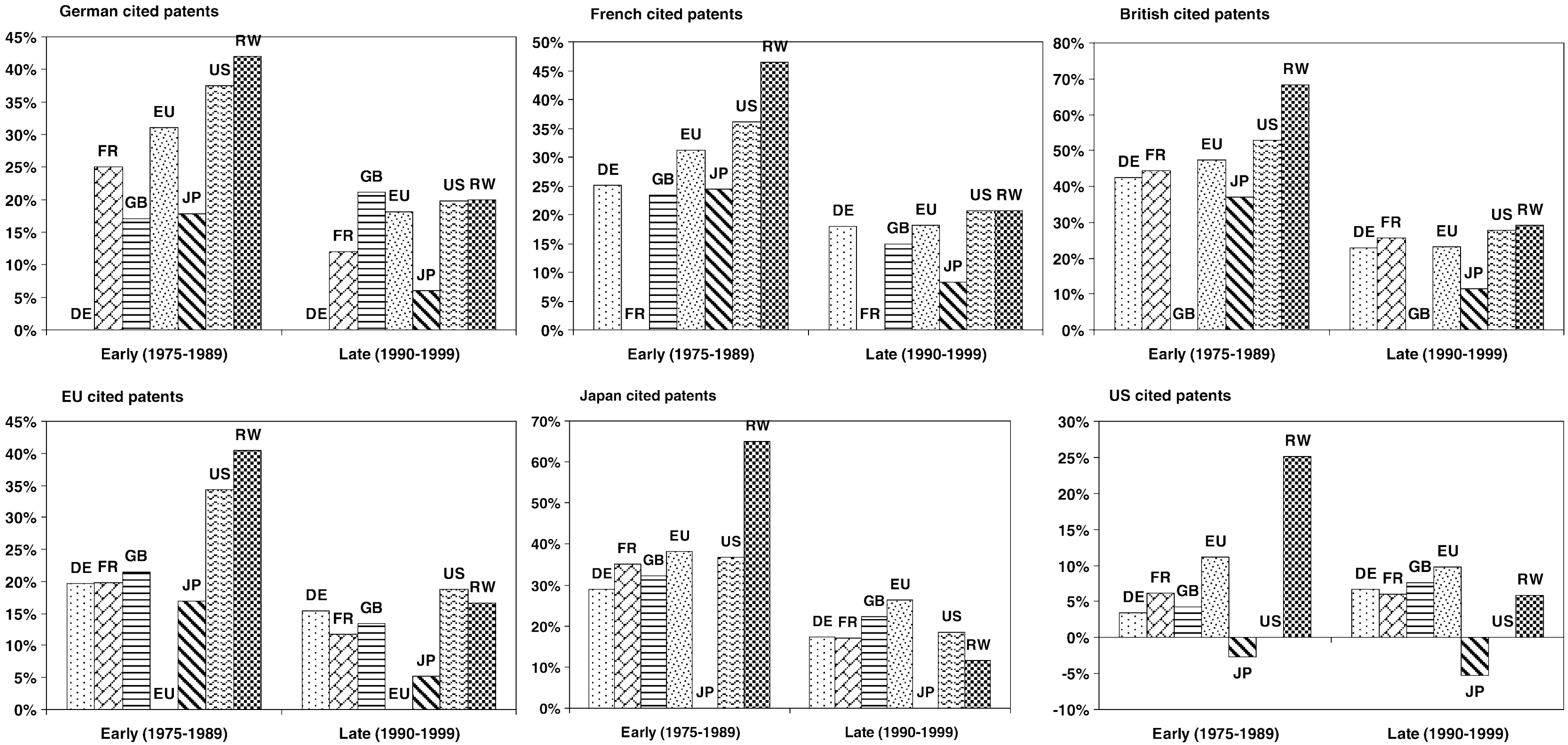

FIGURE 1. Time to first citation, by cited and citing inventor location. This graph shows the relative time (in mean number of day) from the date that a German inventor was granted a patent until the first citation of that patent, by the location of the inventor who made the first citation. For example, the first bar (diagonal bricks) for France in the early period indicates that when the first citation to a Germany patent was made by a French inventor, this citation took on average 25% longer than when the first citation was made by a German inventor (i.e., the mean citation length to a German inventor was 1383 days compared to 1729 days ($1729 = 1383 * 1.25$) to a French inventor). Table 1 shows the raw numbers for all cells. DE = Germany, FR = France, GB = Great Britain, EU = remaining EU countries together, JP = Japan, US = United States, and RW = the rest of the world. In particular, EU consists of Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and Sweden. The abbreviations defined here are used throughout the figures and tables.

Table 1. Time to first citation, by cited and citing inventor location.[a]

Cited	Citing DE	FR	GB	EU	JP	US	RW
				Early Period 1975–1989			
DE	1383	1729	1620	1812	1629	1901	1963
FR	1723	1377	1698	1806	1712	1874	2016
GB	1743	1767	1223	1802	1675	1868	2059
EU	1748	1750	1773	1460	1708	1960	2051
JP	1445	1516	1482	1548	1121	1534	1851
US	1801	1849	1815	1936	1695	1742	2179
RW	1859	1880	1931	1962	1859	2076	1635
				Late Period 1990–1999			
DE	880	986	1066	1040	933	1054	1056
FR	1028	872	1002	1030	944	1052	1052
GB	983	1005	800	985	892	1022	1033
EU	1009	977	991	874	919	1038	1019
JP	897	895	934	965	764	905	853
US	951	945	959	978	844	891	943
RW	999	978	1024	994	851	1014	800

[a]The table shows the mean number of day from the date that a cited inventor was granted a patent until the first citation of that patent, by the location of the inventor who made the first citation. For example, the number in the top panel for the first French (FR) citation to a German (DE) patent in the early period indicates that when the first citation to a Germany patent was made by a French inventor, this citation took on average 1729 days. The top and bottom panels show the average time to first citation for the periods 1975–1989 and 1990–1999, respectively. DE = Germany, FR = France, GB = Great Britain, EU = remaining EU countries together, JP = Japan, US = United States, and RW = the rest of the world. In particular, EU consists of Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and Sweden. The abbreviations defined here are used throughout the tables and figures.

underlying data: the average time to the first citation in the early period from a German inventor to another German inventor was 1383 days compared to 1901 days for an American inventor. This shows that home bias exists. The speed of transmission within Germany increased over time: in the later period, the average time to first citation was only 880 days. But the fall was even greater elsewhere: the time to first American citation fell to 1054.

Looking across Figure 1 as a whole, the pattern is repeated in most regions: foreigners became relatively much quicker at citing domestic patents after 1990.[1] There are, of course, many reasons why the simple patterns in the raw data might be misleading, and much of this paper is devoted to developing and implementing the appropriate econometric tools to show that the results in the raw data are essentially robust to controlling for confounding factors such as unobserved fixed effects and censoring.

[1]There are other interesting features in Figure 1 over and above the general fall in home bias. First, Japanese inventors appear particularly quick to cite other countries' inventors and this trend has grown stronger over time. Second, although home bias has fallen for the United States with respect to the rest of world, it has, if anything, increased with respect to the main EU countries (Germany, France, and Britain). As we see in the econometric section, once we control for other factors, there is not much evidence for home bias of U.S. inventions in the later period.

In terms of related literature, it is well known that tracking international knowledge spillovers is a difficult task. One branch of the literature tries to identify the transfer of technology indirectly by examining changing rates of total factor productivity (TFP) growth across countries and assuming that the faster productivity growth rates of (some) countries or industries that lie further behind the frontier is due to the transfer of ideas.[2] While attractive in its simplicity, a drawback of this approach is that it only provides indirect evidence; the positive correlation between productivity growth and the lagged productivity gap could represent many statistical and economic mechanisms that have nothing to do with the spread of ideas.

A second branch of the literature takes a production function and includes the R&D of *other countries* as an additional variable. These papers tend to find that the R&D of other countries is valuable, but usually not as valuable as R&D in the domestic economy.[3] Most closely related to our current paper, Keller (2002) took this approach and found evidence that technology has become more global over time. This approach has the advantage of using a direct measure of technology; however, it is necessary to identify the relevant external pool of information (i.e., to find a way to appropriately weight the R&D of other countries by order of importance), and the correlation of productivity with R&D is still a very indirect measure of the spillover itself.

A third branch is based around using patent citation information as a direct measure of the transfer of knowledge. The citation of one patent by another strongly suggests that the first patent contained useful knowledge that helped the second innovation. A classic paper in this field is Jaffe, Trajtenberg, and Henderson (1993), who used a matching methodology to show that inventors were far more likely to cite other inventors living in geographic proximity (e.g., the same state or country) compared to inventors in other states or countries. Several papers have followed this approach, and a consensus has emerged that knowledge is subject to a significant degree of "home bias." As with the R&D production function, distance appears to matter.[4] Most closely related to our work, Thompson (2006) used citation data to investigate the localization of knowledge spillovers and found evidence to suggest that this has declined over time.

In this paper, we also use citations to proxy knowledge spillovers, but take a somewhat different approach. We consider the *speed* with which a patent is cited and propose a duration modelling framework that explicitly deals with the problem of unobserved patent characteristics that may be correlated with location or other characteristics. To see how fixed effects could generate a bias, consider the case of two countries—the

[2]For example, see Griffith, Redding, and Van Reenen (2004).

[3]For an introduction to spillovers in general, see Griliches (1992); at the cross-country level, see Coe and Helpman (1995) and Keller (1998); at the industry level, see Bernstein and Mohnen (1998). Work at the firm level finds evidence that countries behind the frontier benefit much more from frontier R&D than vice versa; see Branstetter (2001), Branstetter and Sakakibara (2002), and Griffith, Harrison, and Van Reenen (2006).

[4]For example, Jaffe and Trajtenberg (1999) found that inventors in one country were far more likely to cite inventors living in the same country than in other countries, although this difference tended to diminish over time. Thompson and Fox-Kean (2005) argued that using more disaggregated patent classes drives away localization effects within the United States, but they still observed home bias between the United States and other countries. See also Henderson, Jaffe, and Trajtenberg (2005) for a rejoinder.

United States and Japan. Assume that higher quality patents will be cited more quickly than lower quality patents. If U.S. inventors produce higher quality patents *and* inventors who produce higher quality patents are also more nimble at using the ideas of other countries, then we will observe that U.S. inventors tend to cite other U.S. inventors more than they cite Japanese inventors. This will give the impression of home bias, whereas, in fact, it has to do with the higher average quality of U.S. inventors, which leads to both the generation of new knowledge and the faster absorption of older knowledge. Controlling for fixed effects will, therefore, reduce the degree of home bias observed in naive estimators.[5]

Using a duration model without fixed effects, we find evidence of large home bias, in line with most of the existing literature, but we find that home bias is partly a statistical artefact of the failure to control for unobserved heterogeneity (e.g., differences in patent quality). This heterogeneity has been found to be an important feature of patent values (e.g., Pakes (1986)). Our most important finding is that even after controlling for fixed effects, other covariates and censoring intercountry home bias appears to have *fallen* over time. This observation is consistent with the raw data shown in Figure 1 and Table 1. Other econometric evidence that we are aware of that shows that geography matters less over time is Keller (2002) and Thompson (2006), mentioned above, and Kim, Morse, and Zingales (2006), who found the lower apparent degree of spillovers within elite U.S. university departments.[6] Our work provides new evidence that the geographical localization of knowledge spillovers has fallen over time. Furthermore, the fall in home bias has been greater in the more "traditional" sectors (such as chemicals and mechanical engineering) than the more "modern" technological sectors (such as the information and communication technologies sector and the pharmaceutical sector). This is consistent with the evidence for agglomeration and clustering in these high-tech sectors as suggested by some economic geography models (see Redding (2009)).

Our econometric method builds on Chamberlain (1985) and Ridder and Tunali (1999), which are based on a multiple-spell duration model that is new in the empirical literature on knowledge spillovers. Our method has several important advantages over previously used methods: first, we focus on the first few citations, for which we believe geography matters most (note that we are careful to show the sensitivity of our results to using the different numbers of citations). Second, we allow for a very general form of patent heterogeneity, thus providing new empirical evidence that is unlikely to be driven by different qualities or unobserved characteristics of cited patents. Third, we correct for the censoring problem, which is that newly granted patents are less likely to

[5]The bias is not easily signed. Consider a second scenario where inventions in Japan remain of lower quality on average than in the United States, but Japanese inventors are faster to absorb old knowledge than their U.S. counterparts. This will make it appear that Japanese inventors cite U.S. inventors a lot and could disguise the existence of home bias. In this case, controlling for fixed effects will remove the bias and *increase* the degree of home bias observed in nonfixed-effects estimators. In summary, the fixed-effects bias could go in either direction, but certainly could be important.

[6]A recent paper by Head, Mayer, and Ries (2007) estimated a gravity model of trade for services. As with goods, they found no evidence of distance mattering less for services as a whole. However, for one important subsector, "miscellaneous business services," distance does appear to matter less in 2004 than in 1992.

be cited by other patents, hence avoiding the standard problem in analyzing patent citation data. The method we apply has a number of potential applications beyond the one we investigate in this paper; for example, one could analyze the degrees of importance of geography within the United Sates across different states. Another possible application is to look at the extent to which knowledge spreads differently across institutions, such as universities, private firms, and government labs.

The paper is laid out as follows. Section 2 sketches our econometric model. Section 3 details the data and Section 4 gives the results. Some concluding comments are given in Section 5. Appendix A provides the details of our estimation method, the asymptotic distribution of our estimator, and comparisons between our econometric model and related models in the literature. Appendix B gives additional data description and estimation results.

2. Modelling strategy

Consider a set of inventions $i = 1, \dots, I$ and a set of inventors $j = 1, \dots, J$. We take this pool of inventions and inventors as exogenously determined,[7] with their numbers growing over time. The inventors will "learn"[8] of invention i after a time period T_{ij}. We think of T_{ij} as the "diffusion lag" between invention i and inventor j.

Time T is a nonnegative random variable with distribution $F(t)$ and density $f(t)$. There are several factors that determine the diffusion lag, including characteristics of the invention Z_i, characteristics of the inventor Z_j, and the joint characteristics of the invention–inventor match Z_{ij}. There is a set of nongeographical variables that influences the speed at which information flows. For example, news of a higher quality invention may travel more quickly as will inventions in more established technological fields compared to newer areas. Similarly, lower quality inventors may be slow to pick up on news of new technologies. Finally, information will diffuse more quickly for inventors and inventions operating in the same technological field compared to those operating in different fields.

Our main interest is in geographical barriers to knowledge transfusion as proxied national boundaries. Thus, we hypothesize that the nongeographical factors determine the expected diffusion lag, but there will be an additional cost of transmitting information depending on whether inventors are located in the same country as an invention or are in a different country. To the extent that this slows down the diffusion of knowledge, we will say that there is a home bias. Note that this home bias exists over and above any effect that arises from inventions or inventors being intrinsically faster (or slower) in picking up knowledge in general. We can control for these effects by linear country dummies of invention (CTY_i) and inventor (CTY_j), with the key home bias term being whether the particular pair of countries (CTY_{ij}) matter for diffusion.

[7] Many general equilibrium growth models seek to derive the stocks of inventions and inventors as endogeneously related to the diffusion lag (e.g., Cabellero and Jaffe (1993)), but we abstract from these considerations here.

[8] Learning can be interpreted in different ways. It is a combination of becoming *aware* of the invention, *understanding* it, and then finding it *useful* enough to build on to develop new knowledge.

The hazard function of the diffusion lag is defined as $\frac{f(t)}{1-F(t)}$, which we model as a function of observables X_{ij} that incorporate the empirically observable counterparts to Z_{ij} and Z_j, and an unobservable fixed effect U_i, which absorbs all the factors specific to the cited patent (such as quality).

In our application, inventions are measured by cited patents, inventors are measured by citing patents, and the diffusion lag is measured by the duration of the citation lag between invention and inventor. We estimate the impact of home bias on knowledge spillovers using a multiple-spell duration model. Consider a patent that is taken out (the *cited* patent) and the patents that subsequently cite it (the *citing* patents). If geography is important for the flow of information, then we should expect to see that durations are shorter when the *citing* inventor is located near the *cited* inventor. We focus on the first few citations. Geography matters because most of the knowledge in a new invention is tacit, whereas over time, this information becomes codified. Consequently, over time, information about the invention is more easily transmitted across greater distances, and researchers with direct knowledge of the invention become more geographically dispersed. We see evidence of this in the raw patents data. For example, when we look at all patents taken out by German firms and we look at who first cites those patents, in 17% of cases it is another firm located in Germany; when we look at the fifth time the patent is cited, then 12% are firms located in Germany; by the tenth time the patent is cited, 10% are firms located in Germany. Looking across other locations, we see that the share of cases where the cited and citing firm are in the same country falls monotonically, with a higher share of the first citations being in the same country.

As highlighted above, unobserved heterogeneity could confound our estimates, as higher quality patents may be cited more quickly. To control for this, we use an estimator that is analogous to the linear difference estimator by comparing the first and second citations for each cited patent. By comparing the difference between the citing patents, we are able to "difference out" the unobserved characteristics of the cited patent.[9]

Let subscript i index cited patents and let subscript j index citing patents. Under this convention, let Y_{ij}^* denote the jth citation duration for the ith patent, that is, the number of days from the date when the ith cited patent is granted to the date when the jth citing patent is granted, where $i = 1, \dots, n$ and $j = 1, \dots, J$.[10] Here n is the number of patents and J is the number of (potential) citations for each cited patent. Also, let X_{ij} denote the attributes of the jth citing patent for the ith cited patent and let U_i denote unobserved characteristics of the cited patent. For example, U_i may represent unobserved quality of the cited patent.

We consider a multiple-spell version of the mixed proportional hazards model. The hazard that $Y_{ij}^* = y_{ij}^*$ conditional on $X_{ij} = x_{ij}$ and $U_i = u_i$ has the form

$$\lambda_i(y_{ij}^*) \exp(x_{ij}'\beta + u_i), \tag{1}$$

[9]See, for example, Chamberlain (1985), Ridder and Tunali (1999), Horowitz and Lee (2004), and Lee (2008).

[10]The notation Y_{ij}^* is used to reflect that Y_{ij}^* is a latent variable due to the usual right censoring problem. In fact, we observe $Y_{ij} = \min(Y_{ij}^*, C_i)$ and $\Delta_{ij} = 1(Y_{ij}^* < C_i)$, where $1(\cdot)$ is the usual indicator function and C_i denotes the censoring time. See Appendix A for details on how to handle censoring.

where β is a vector of unknown parameters and $\lambda_i(\cdot)$ is a cited-patent-specific baseline hazard function.

The citation durations Y_{ij}^* are assumed to be independent of each other, conditional on the observed and unobserved characteristics (X_{ij}, U_i). In addition, the observed covariates X_{ij} are assumed to be constant within each spell but to vary over spells. For example, X_{ij} may include the location of the inventor of the jth citing patent for the ith cited patent. We allow U_i to be arbitrarily correlated with X_{ij} and do not impose any distributional assumptions on U_i; therefore, U_i is a *fixed effect*. The multiple-spell structure allows U_i to have a very general form, compared to unobserved heterogeneity in the single-spell duration models. The functional form of the baseline hazard function $\lambda_i(\cdot)$ is unspecified and it can also vary across different cited patents. Therefore, the model also allows for unobserved heterogeneity in the shape of the hazard function.[11]

The conditional independence assumption is indispensable in our econometric modelling strategy. The important implication of this assumption is that it requires that one citation does not *lead* to another citation. What would cause us problems is if the first citation of a patent provided information to other potential citers and, therefore, affected the duration to the next citation.[12]

Under the conditional independence assumption, such that Y_{ij}^* are independent of each other conditional on (X_{ij}, U_i), we can estimate β using a conditional likelihood approach (e.g., Chamberlain (1985), Ridder and Tunali (1999)). The idea behind the conditional likelihood approach is as follows. Assume that there are only two potential citing patents ($J = 2$). The probability that the observed first citation duration is first, conditional on the duration of the first citation, is given by

$$\begin{aligned}
&\Pr\bigl[Y_{i1}^* \leq Y_{i2}^* | \min\{Y_{i1}^*, Y_{i2}^*\} = y_{1i}^*, X_{i1} = x_{i1}, X_2 = x_{i2}, U_i = u_i\bigr] \\
&\quad = \frac{\lambda_i(y_{1i}^*)\exp(x_{i1}'\beta + u_i)}{\lambda_i(y_{1i}^*)\exp(x_{i1}'\beta + u_i) + \lambda_i(y_{1i}^*)\exp(x_{i2}'\beta + u_i)} \\
&\quad = \frac{\exp(x_{i1}'\beta)}{\exp(x_{i1}'\beta) + \exp(x_{i2}'\beta)},
\end{aligned} \tag{2}$$

which does not depend on u_i or λ_i. Therefore, β can be estimated based on this conditional likelihood without the "incidental parameters" problem.[13]

[11]The heterogeneity term U_i is not separately identified from the baseline hazard function $\lambda_i(\cdot)$. The model in (1) can be rewritten as $\tilde{\lambda}_i(y_{ij}^*)\exp(x_{ij}'\beta)$ with $\tilde{\lambda}_i = \exp(u_i)\lambda_i$.

[12]While this is of course possible, we believe that it is not a major problem in our context because we are focussing on first and second (or third and fourth) citations. Due to the publication lag, the first citation is often not public by the time the second citation is made.

[13]Thompson (2006) uses Chamberlain's (1980) conditional logit model to estimate the effects of localized knowledge spillovers. His paper is different from ours in two main ways. First, he used pairs of citing and cited patents to construct the binary matching indicators (the dependent variable), whereas we start from a multispell duration model and then used only the first few citations. Second, Thompson (2006) used an interaction term between the indicator variable for inventor citations and the cited patent age to identify the effects of knowledge spillovers, whereas we use the location of the inventor of a citing patent. See Section A.3.2 for more details.

A usual problem with analyzing such data is censoring. Given any data set, there will be some patents that have not (yet) been cited, but that could in the future be cited. The standard conditional likelihood approach (see, e.g., Chamberlain (1985)) can handle censoring if one always observes covariates X_{ij}. In our application, like many others, X_{ij} are only observed when durations are uncensored. For example, we can identify the location of the inventor of a citing patent only in the case when it is observed. This problem forces us to use only uncensored spells and this may introduce a selection problem. In our data, citation durations are obtained by looking at all recorded citations at a particular date (December 31, 1999). We therefore treat the censoring as independent of citation durations and covariates (what we need is that the application and grant dates are independent of quality), and then weight the observations by the inverse of the propensity to observe complete spells. This is analogous to the way that missing data are treated in inverse probability weighted estimation (e.g., Wooldridge (2007)). See Appendix A for details of our estimation method.

There are two main differences between our approach and the more usual Jaffe and Trajtenberg (1999) approach. First, a major advantage is that we can control for unobserved heterogeneity in a way that they do not. Consistent with Thompson and Fox-Kean (2005), we find that using three digit technology classes is an inadequate control, as the number of rejections of home bias fall substantially when we include our fixed effects over and above these technology dummies. Second, as with any fixed-effect estimator, a potential disadvantage of our approach is that we use only a subsample of the data that they use (two or more cites instead of all cites). We do not attempt to characterize the entire shape of the citation function, but rather focus on the first few cites. We believe that this is a natural approach to examining international spillovers, as localization effects should be strongest soon after a patent is granted when knowledge is still mostly tacit. Nevertheless, we see this approach as a complement rather than a substitute for the Jaffe and Trajtenberg (1999) model.[14] See Section A.3 for a more detailed comparison of our approach with others in the literature.

3. Data

To implement this estimator, we use data from the National Bureau of Economic Research (NBER) U.S. Patent Citations Data File.[15] These data include information on all patents taken out at the United States Patent Office (USPTO) and have been widely used in the economic analysis of spillovers.

Table 2 shows the sample sizes for our analysis. The NBER data consist of patents granted and citations made to these patents between 1975 and 1999. In total, we use data on over 2.1 million cited patents. While these patents were all taken out in the USPTO, the assignees and inventors can be located anywhere in the world. We use the information on the inventors' addresses to identify the location of the patent.[16] We focus

[14]See Belenzon and Van Reenen (2007) for evidence on the changing time patterns of citations using an approach closer to Jaffe and Trajtenberg (1999).

[15]See Jaffe (1986), Hall, Jaffe, and Trajtenberg (2001, 2005), and Jaffe and Trajtenberg (2002).

[16]Where there is more than one inventor, we follow Jaffe, Trajtenberg, and Henderson (1993) and allocate patents to the country where the majority of inventors are located. In the case of ties, we randomly choose one of the countries.

Table 2. Sample sizes of patent citation data.[a]

Technological Category	Period	DE	FR	GB	EU	JP	US	RW	Total
		Country of Cited Patents							
Chemical	All	46,697	13,840	14,414	21,662	73,211	231,594	27,714	429,132
	Early	26,663	7355	8802	11,173	32,385	130,532	14,388	231,298
	Late	20,034	6485	5612	10,489	40,826	101,062	13,326	197,834
Computers and communications	All	8485	6725	6236	7781	70,657	134,335	12,830	247,049
	Early	4094	3137	2713	2904	19,808	45,308	2763	80,727
	Late	4391	3588	3523	4877	50,849	89,027	10,067	166,322
Drugs and medical	All	12,578	6992	7862	9887	18,044	115,365	12,612	183,340
	Early	5841	2741	3494	3391	6763	38,777	4472	65,479
	Late	6737	4251	4368	6496	11,281	76,588	8140	117,861
Electrical and electronic	All	25,723	12,029	10,585	13,942	85,591	193,424	25,467	366,761
	Early	14,251	6374	6448	7300	30,747	97,099	8003	170,222
	Late	11,472	5655	4137	6642	54,844	96,325	17,464	196,539
Mechanical	All	46,260	13,976	13,837	24,266	96,811	240,766	31,535	467,451
	Early	26,429	8220	8979	14,009	42,672	133,759	14,822	248,890
	Late	19,831	5756	4858	10,257	54,139	107,007	16,713	218,561
Others	All	30,064	11,452	12,117	21,711	46,330	284,448	38,853	444,975
	Early	17,475	6519	7438	12,214	21,275	151,837	17,383	234,141
	Late	12,589	4933	4679	9497	25,055	132,611	21,470	210,834
Total	All	169,807	65,014	65,051	99,249	390,644	1,199,932	149,011	2,138,708
	Early	94,753	34,346	37,874	50,991	153,650	597,312	61,831	1,030,757
	Late	75,054	30,668	27,177	48,258	236,994	602,620	87,180	1,107,951

[a]Data consist of patents that were granted between 1975 and 1999. The patents in the data were all taken out at the United States Patent Office (USPTO). A country of cited patents refers to the location of an applicant. The period "All" includes years from 1975 to 1999 in which cited patents are granted; "Early" and "Late" periods correspond to 1975–1989 and 1990–1999, respectively.

on inventors located in the G5 countries—the United States, Japan, France, Germany, and Great Britain. We group the remaining EU countries together[17] and then consider the rest of the world (RW) as the residual category. Unsurprisingly, the United States is the leading country with nearly 1.2 million patents, and Japan is second with nearly 400,000. We split our sample into two subperiods, 1975–1989 and 1990–1999, and consider whether the evidence for home bias differs over these two periods.

Crucially for our purposes, the NBER data contain information on all subsequent citations to each patent made by other patents. In our baseline results, we use the information contained in the first and second citations to implement the estimator described in the previous section. As highlighted above, an issue that arises with using citation data is the problem that for some patents (those taken out near the end of the period), these citations will be censored; that is, the first or second citation will not have occurred yet. This is a well documented problem with using citation data.[18] For example, in our data

[17]These are Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and Sweden.

[18]For example, Hall, Jaffe, and Trajtenberg (2001) and Bloom and Van Reenen (2002).

Table 3. Summary statistics for patent citation data.[a]

Variable	Chemical	Computers and Communications	Drugs and Medical	Electrical and Electronic	Mechanical	Others
Proportion of patents						
With two or more citations	0.56	0.62	0.49	0.60	0.54	0.49
With only one citation	0.15	0.12	0.14	0.14	0.17	0.17
With no citation	0.28	0.26	0.37	0.25	0.30	0.34
Proportion of self-citation						
First citation	0.22	0.15	0.19	0.15	0.17	0.14
Second citation	0.19	0.13	0.15	0.13	0.14	0.12
Proportion of same technology class						
First citation	0.65	0.71	0.76	0.67	0.68	0.68
Second citation	0.63	0.71	0.75	0.65	0.67	0.66
Average of base						
First citation	0.20	0.18	0.15	0.11	0.13	0.16
Second citation	0.21	0.19	0.17	0.12	0.14	0.17

[a]Data consist of patents that were granted between 1975 and 1999. The patents in the data were all taken out at the United States Patent Office (USPTO). The base variable is defined as the number of patents in the citing country and technology subcategory for the citing year (1 unit = 1000 patents).

(see Table 3), for 28% of patents in the chemical technology sector, we never see a citation, for 15% we see only one citation, and for the remaining 56% we see two or more citations. Similar patterns are observed for other technology sectors. Because of this, it is important that our empirical methods correct for censoring biases.

We control for whether the citation is a self-citation (i.e., whether the assignee is the same on the cited and citing patent) and whether the cited and citing patent are in the same technology class. We also control for the size of the base of potential citing patents, that is, the number of patents in the citing country and technology subcategory for the citing year. We discuss the interpretation of this variable in Section 4.4 below; we believe it reflects the explosion of patenting that led to some diminution of average patent quality (see Kortum and Lerner (1999), Jaffe and Lerner (2004)).

Table 3 reports some summary statistics for these control variables by technology category. In chemicals, 22% of all first citations are self-citations; this falls to 19% for the second citation. On average across technology sectors, just under 20% are self-citations and this value declines by 2%–4% from the first citation to the second citation. More than 65% of citations are from the same technology class. The proportions of self-citations, same technology class, and the averages of the bases (potential cites) are characteristics of citing patents and thus they are obtained from only complete citation spells.

4. Results

4.1 *Basic results*

We implement the estimator described above on all patents granted by the USPTO between 1975 and 1999. We report results across seven regions and six technology categories, and allow all the coefficients to vary across these groups.

Table 4. Estimation results for the chemicals technology category. The country of cited patents is Germany (DE). The sample size is 46,697. There are 25,016 observations with at least two citations.[a]

Variable	No Fixed Effect (1)	Fixed Effect (2)	Fixed Effect Plus Censoring (3)
FR	−0.15	−0.03	−0.03
	(0.04)	(0.07)	(0.08)
GB	−0.03	0.03	0.02
	(0.04)	(0.06)	(0.08)
EU	−0.12	−0.04	0.04
	(0.03)	(0.05)	(0.07)
JP	0.03	0.00	0.00
	(0.02)	(0.04)	(0.05)
US	−0.08	−0.02	0.05
	(0.02)	(0.04)	(0.05)
RW	−0.13	−0.12	−0.14
	(0.03)	(0.05)	(0.07)
Self-cit.	0.38	0.39	0.48
	(0.02)	(0.04)	(0.04)
Tech. class	0.16	0.15	0.15
	(0.01)	(0.02)	(0.03)
Base	−0.15	−0.53	−0.80
	(0.06)	(0.09)	(0.12)

[a] Standard errors are given in the parentheses. The dummy variables for the location of an applicant of citing patent correspond to Table 1. The omitted category in citing patent country dummies is Great Britain (GB). The self-citation and technology class variables are dummy variables. The base variable is the number of patents in citing country and subcategory for the citing year (1 unit = 1000 patents). Different columns show different estimates. Column (1) shows no-fixed-effect estimates using the only the first citation duration, column (2) shows fixed-effect (FE) estimates using the first two citation durations, and column (3) shows FE estimates with censoring taken into account.

4.1.1 *An example—chemical engineering* We begin by going through the results for one technology category in one country to illustrate our methodology. In Table 4, we show the coefficient estimates for the citing country dummies when we look at chemical engineering in Germany. Each column in Table 4 reports the results from a different regression. The omitted category is own country—the location of the cited patent—which in this case is Germany (DE). There are potentially 46,697 cited patents in chemical engineering in Germany over this time period; from this sample, 25,016 patents are cited at least twice. The main variables of interest are the indicators of the country of the citing firm. Also included in the regression is an indicator of whether the citation is a self-citation, whether the cited and citing patent are in the same technology class (three digit), and the total number of citing patents in that country and technology class for the citing year.

In column (1) of Table 4, we estimate the coefficients using a proportional hazard model with only the first citation duration. This is equivalent to our model without fixed effects (and constraining the baseline hazard to be the same across patents), that is, compared to equation (1), we assume

$$\lambda(y_{ij}^*)\exp(x_{ij}'\beta). \tag{3}$$

To keep the sample the same as when we estimated the fixed-effects model, we restrict the estimation to patents with at least two citations. The coefficients on the country dummies indicate whether inventors located in that country cite German inventors in chemical engineering faster (a positive coefficient) or slower (a negative coefficient) than inventors from the omitted category (which is always own country, in this case Germany). If there is home bias, we expect *negative* coefficients on the other country dummies, that is, they are slower to cite than home inventors. In column (1), we see negative and significant coefficients on four country dummies; these suggest strong support for home bias. Japanese inventors are the swiftest foreign group to cite German inventors: they are actually 3% faster than German inventors themselves, although the estimated coefficient is insignificant at the 5% level. By contrast, inventors in France are 15% slower to cite German patents.

In column (2) of Table 4, we control for unobserved cited patent characteristics (e.g., quality), which may be correlated with the speed with which the patent is cited, by estimating the coefficients using the fixed-effect estimator (without correcting for censoring).[19] When fixed effects are included, most coefficients become closer to zero and all country dummies become statistically insignificantly different from zero, except for the rest of the world (RW). This suggests that failure to control for unobserved heterogeneity increases the degree of home bias.[20] The simple fixed-effects estimator in column (2) ignores the problem of censoring. In column (3), we also allow for censoring, which leads to little change in most of the coefficients (but increases the standard errors a bit) and has relatively little effect on the qualitative findings. As would be expected, if the patent is taken out by the same assignee (a self-citation), the citation speed is significantly faster (about 48% faster than non-self-citations in column (3)). Similarly, patents in the same technology class cite each other significantly faster (15% faster than patents in different technology classes according to column (3)). Patents in larger country–technology classes are cited less frequently.

We continue to illustrate the method by looking across all countries, but still restricting ourselves to patents in the chemical engineering category. In Table 5, each row contains parameter estimates from a separate multiple-spell duration model for each country. For example, the first row shows the results from column (3) of Table 4 (the coefficients on self-citation, technology class, and base are not reported). Table 5 shows only the results for the fixed effects and censoring model (denoted FE + C), that is, the model shown in column (3) of Table 4.

What do the coefficients in Table 5 tell us? As before, the omitted base category is always the home country, and negative coefficients suggest home bias. Looking across the second row for France, we see that only inventors from the rest of the world (mainly developing countries) are significantly slower to cite French inventors than the French themselves: the coefficients for German, British, EU, Japanese, and U.S. inventors are

[19]Specifically, the estimator maximizes the likelihood equation (A.1) in Appendix A without the correction term $G_n(\max[Y_{i1}, Y_{i2}])$.

[20]It is possible to have a case in which failure to control for unobserved heterogeneity decreases the degree of home bias, since the direction of bias from failure to control for fixed effects cannot be signed a priori.

TABLE 5. Estimation results of chemical (FE + C): full sample.[a]

	Country of Citing Patents						
Country of Cited Patents	DE (1)	FR (2)	GB (3)	EU (4)	JP (5)	US (6)	RW (7)
DE		−0.03	0.02	0.04	0.00	0.05	−0.14
		(0.08)	(0.08)	(0.07)	(0.05)	(0.05)	(0.07)
FR	0.03		−0.01	−0.25	−0.10	−0.02	−0.33
	(0.11)		(0.17)	(0.20)	(0.12)	(0.12)	(0.14)
GB	−0.10	−0.32		0.00	−0.06	−0.05	−0.29
	(0.09)	(0.13)		(0.13)	(0.09)	(0.09)	(0.11)
EU	0.07	−0.06	0.13		−0.06	−0.04	−0.18
	(0.08)	(0.12)	(0.13)		(0.09)	(0.09)	(0.10)
JP	−0.03	−0.23	−0.17	−0.24		0.03	−0.24
	(0.04)	(0.08)	(0.07)	(0.07)		(0.03)	(0.05)
US	0.01	−0.15	−0.10	−0.22	−0.06		−0.24
	(0.02)	(0.04)	(0.04)	(0.03)	(0.02)		(0.03)
RW	0.15	−0.14	0.01	−0.12	0.03	0.08	
	(0.09)	(0.16)	(0.12)	(0.12)	(0.08)	(0.08)	

[a]Each row contains parameter estimates and their standard errors (in parentheses) from a separate multiple-spell duration model for each country. The censored fixed-effect estimator (FE + C) is used with the entire sample for the technology category mechanical. The country name in the first column corresponds to the location of the patent's inventor, which is subsequently cited. The country names in columns (1)–(7) correspond to the inventor location of the patent that subsequently cites the original patent. The omitted base country dummy is the cited patent's country. In addition to country dummies, each hazard regression includes, as explanatory variables, dummy variables for self-citation and technology class, and the number of patents in citing country and subcategory for the citing year.

insignificant, and inventors from the rest of the world are 33% slower to cite French inventors. So, just as in the German case, we do not see home bias after controlling for unobserved heterogeneity within the main developed nations. Note that all regressions include unreported controls for whether the citation is a self-citation, whether it is in the same technology subcategory (three digit), and the total number of citing patents in that country and technology class (base). Most of these controls are highly significant and lead to the impression of home bias if omitted.[21] The story is different if we look at the United States (the sixth row in Table 5). All countries except Germany are significantly slower to cite U.S. inventors than the Americans themselves: the French inventors are 15% slower to cite U.S. inventors, British are 10% slower, other Europeans are 22% slower, Japanese are 6% slower, and the rest of world is 24% slower. A similar pattern exists for Japan: the European countries are much slower to cite Japanese patents than the Japanese themselves.

We give a graphical representation of the results from Table 5 in Figure 2 to make it easier to eyeball the results. Each cell corresponds to the equivalent cell in Table 5. A circle represents a negative coefficient (home bias) and a cross represents a positive coefficient. The size of the circle or the cross corresponds to the level of statistical significance of a one-sided test for the null hypothesis that the corresponding coefficient is zero. A large circle represents significance at the 1% level, a medium circle represents

[21]This is true for all econometric models, as seen in all columns in Table 4.

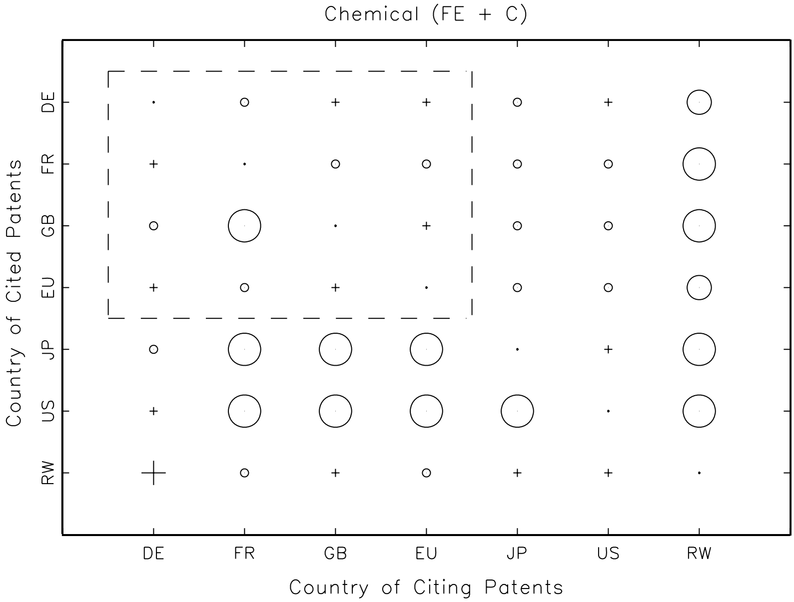

Figure 2. Graphical representation of estimation results. Each cell corresponds to the equivalent coefficient in column (3) of Table 4. A circle represents a negative coefficient (home bias) and a cross represents a positive coefficient. The size of the circle or the cross corresponds to the level of statistical significance of a one-sided test for the null hypothesis that the corresponding coefficient is zero. A large circle represents significance at the 1% level, a medium circle represents significance at the 5% level, a small circle represents significance at the 10% level, and a tiny circle represents *insignificance* at the 10% level. The same ordering applies to crosses. The leading diagonal corresponds to the omitted variable in each regression and, therefore, no coefficient is estimated. The upper left quadrant with dashed lines contains the cross-citations from the European countries.

significance at the 5% level, a small circle represents significance at the 10% level, and a tiny circle represents *insignificance* at the 10% level. The same ordering applies to crosses. The leading diagonal corresponds to the omitted variable in each regression; therefore, no coefficient is estimated or displayed.[22] So it is possible to immediately detect the degree of home bias for a country by looking at the number and size of circles across a row. The United States as country of cited patents, for example, has a full row of large circles, indicating significant home bias, whereas European countries do not (this feature is not apparent from the raw data from all sectors in Figure 1). It is also clear from Figure 2 that there is less home bias among the EU countries (points in the top left quadrant marked with the dashed box), compared to between the non-EU countries and EU countries. The top right quadrant contains no rejections for the Japan and U.S. columns, suggesting that Japan and the United States are no slower in citing European patents than Europe's own inventors; however, the bottom left quadrant contains many rejections for the Japan and U.S. rows, suggesting that European countries (except Ger-

[22] A full set of results are available on request from the authors.

Table 6. Number of rejections of no home bias using entire sample.[a]

Technological Category	Max. No. of Rejections	No FE 10% (1)	No FE 5% (2)	No FE 1% (3)	FE 10% (4)	FE 5% (5)	FE 1% (6)	FE + C 10% (7)	FE + C 5% (8)	FE + C 1% (9)
Chemical	42	34	32	23	16	12	9	14	14	12
Computers and communications	42	18	17	10	16	14	14	17	15	15
Drugs and medical	42	26	24	18	15	11	5	15	10	6
Electrical and electronic	42	22	19	16	13	13	11	15	14	12
Mechanical	42	26	25	17	16	13	6	15	11	9
Others	42	36	33	30	20	14	11	15	13	11
Total	252	162	150	114	96	77	56	91	77	65
Percentage		0.64	0.60	0.45	0.38	0.31	0.22	0.36	0.31	0.26

[a] The number of rejections of one-sided t-tests for individual coefficients is shown in each cell of the table. Three levels of tests are considered: 1%, 5%, and 10%. Also, three different estimators are used: no-fixed-effect estimator (no FE) using only the first citation duration, fixed-effect (FE) estimator using the first two spells, and censored fixed-effect (FE + C) estimator.

many) are slower in citing Japanese and U.S. patents. Hence, there exists an interesting asymmetry between the European block and the Japan/U.S. block, in the sense that European inventors are slow to cite Japanese and American patents, but inventors located in Japan and the United States are quick to cite European patents. Another interesting asymmetry exists: the rest of the world is slow to cite the main developed countries, while the main developed countries are quick to cite the rest of the world.[23]

4.1.2 *Main results* We conduct the equivalent analysis across all seven regions and six sectors. Table 6 summarizes the results (full results available on request). The number of rejections of one-sided t-tests for the coefficients on country dummies are shown for each sector. Test results are shown for three levels (1%, 5%, and 10%) using the no-fixed-effect hazard model estimator (No FE), the fixed-effect estimator (FE), and the censored fixed-effect estimator (FE + C).

The first striking result in Table 6 is that there appears to be strong evidence for home bias when we consider the model that does not control for unobserved heterogeneity (columns (1)–(3)). Of the 252 tests[24] for no home bias, we reject 150 at the 5% level, or around 60%. This is consistent with evidence from the analysis of citations data in other econometric studies (Jaffe and Trajtenberg (1999), Henderson, Jaffe, and Trajtenberg (2005), Thompson and Fox-Kean (2005)). However, the picture changes when

[23] Germany is different from the other European countries in that it is particularly quick to cite other countries in the category chemicals, but not other industries (see Figure 3). This may be because Germany has a long-standing comparative advantage in the chemical industry. Arora, Landau, and Rosenberg (1999) emphasized the historically strong international links of scientists working in organic chemistry in Germany. Another possible reason is that public sector investment in applied research in Germany has taken quite a different form than in other countries, notably major investments by the government in the Fraunhofer Institutes, which include several located in the United States (see http://www.fraunhofer.de/en/).

[24] Seven country regressions and six country dummies for each regression gives 42 tests for each sector.

we control for unobserved heterogeneity (columns (4)–(6)). Comparing column (5) to column (2), for example, the rejection rate (the proportion of possible rejections that are in fact rejected) falls from 60% to 31%. In other words, there are far fewer rejections of home bias once we control for unobserved heterogeneity. Controlling for censoring makes relatively little difference to the total number of rejects in columns (7)–(9), the rejection rate is the same in column (8), where we control for censoring, as in column (5), where we do not, although it does affect some of the individual results.

The impact of controlling for unobserved cited patent effects can also be seen graphically in Figure 3. For each sector, the left-hand side diagrams shows the pattern *without* controlling for fixed effects (no FE), whereas the right-hand side presents results from our preferred specifications with controls for fixed effects and censoring (FE + C). It is clear that the proportion of large circles (evidence of significantly slower citations by another country) falls when moving from the no-fixed-effects specifications to the preferred specifications. This phenomenon is much less apparent in computers and communications: in Table 6 the number of rejections generally halves when we move between the no-fixed-effects specifications of column (2) to the fixed-effects specifications of column (5), yet for computers and communications the number of rejections essentially remains the same. There are far fewer rejections even without fixed effects for computers, while once we control for unobserved quality (column (5) of Table 6), the number of rejections is quite similar across industries.

Why does unobserved heterogeneity not lead to the same sort of bias in computers as it did in other industries? That is hard to say. The bias from omitted unobserved heterogeneity is not easily signed and could, in principle, go in either direction (see footnote 5). In the raw data (when we do not control for unobserved quality), it seems that the computer industry is very international compared to other industries, yet when we control for unobserved quality, this wipes out most of this difference by reducing the evidence for home bias in other industries.

A second feature of Table 6 and Figure 3 is that the models without fixed effects suggest a sectoral pattern with less home bias in the modern sectors of electrical and electronic and computers and communications than in the more traditional sectors (e.g., chemicals). This is similar to Peri (2005), who found that knowledge spreads much more quickly across regional boundaries in the computer and communication sector. However, once we control for unobserved heterogeneity, the sectors look relatively similar.

A third feature of Figure 3 is that the rest of world (mainly non-OECD (Organization for Economic Cooperation and Development) countries) is consistently slower in citing the patents of the OECD countries. This suggests that non-OECD countries are more "cut off" from international pools of knowledge, because of either their distance, their infrastructure, or their development levels.

4.2 *Falling home bias over time?*

We now turn to the important issue of whether home bias has fallen over time, as some commentators have suggested (e.g., due to the falling costs of international communication and/or travel). We divide our sample into an early period (1975–1989) and a late

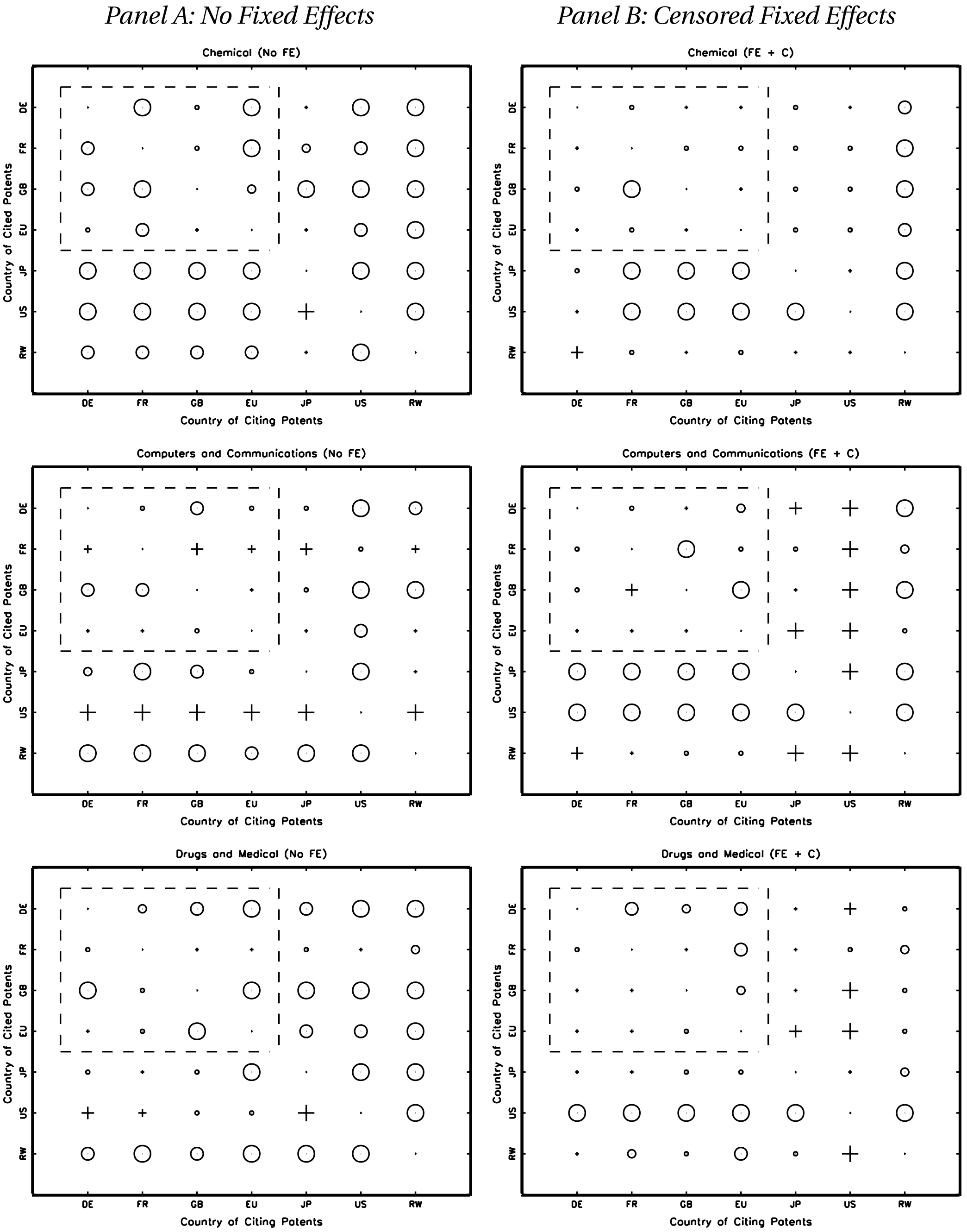

FIGURE 3. No fixed effects (No FE) and fixed effects with censoring (FE + C). For each sector, the left-hand side diagram shows the pattern without controlling for fixed effects, whereas the right-hand side presents results from our preferred specifications with controls for fixed effects and censoring. The upper left quadrants with dashed boxes contain the cross-citations from the European countries.

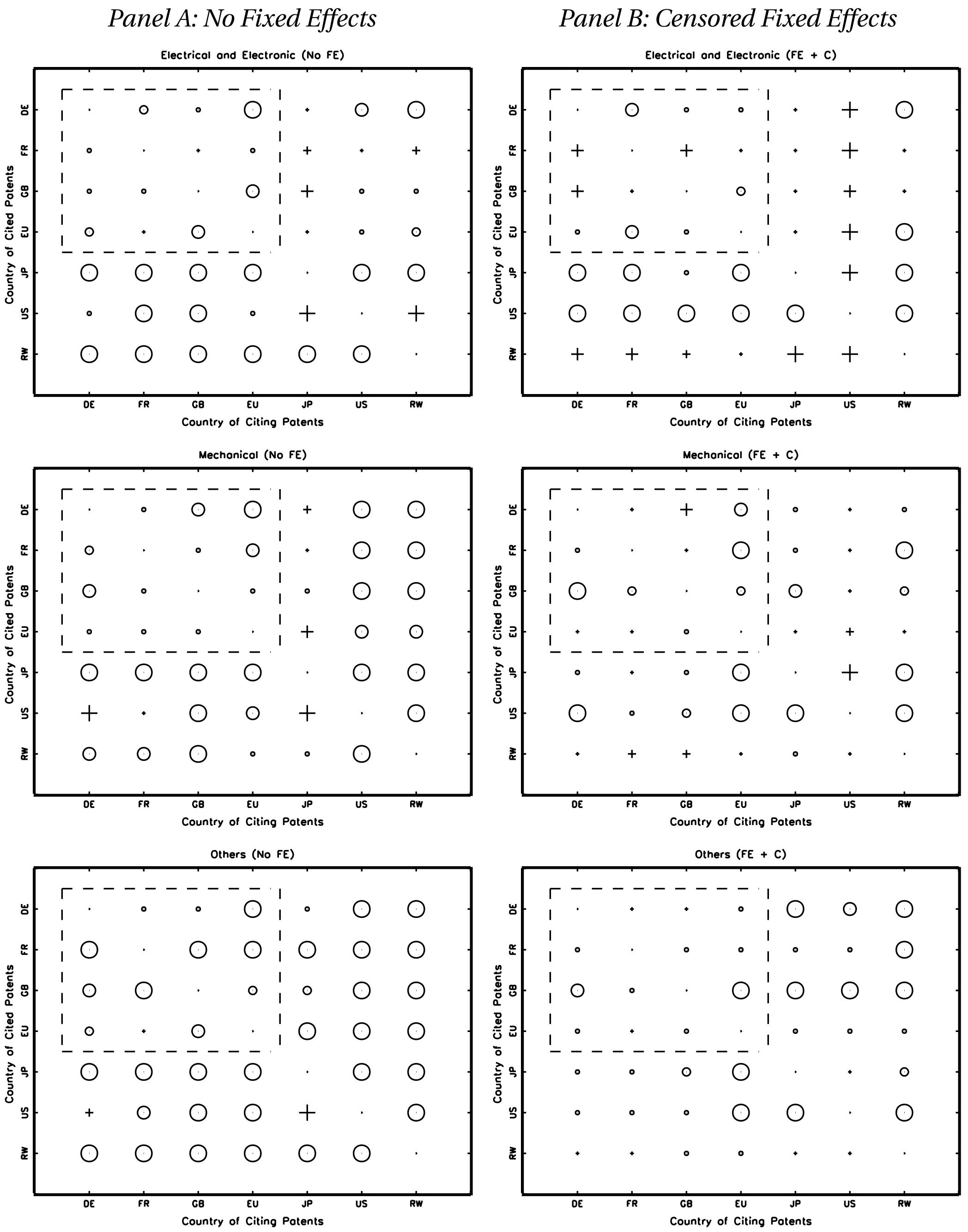

FIGURE 3. *Continued.*

period (1990–1999), where there are a similar absolute number of citations in each period (see Table 2). We reestimate all of our models on these two subperiods separately. We report a summary of these results in Tables 7 and 8 and Figure 4.[25] It is particularly

[25]The full results of these estimations are available on request from the authors.

TABLE 7. Number of rejections of no home bias using subsamples. The estimation method is the no-fixed-effect estimator.[a]

Technological Category	All Countries Early (1)	All Countries Late (2)	OECD Countries Early (3)	OECD Countries Late (4)	EU Countries Early (5)	EU Countries Late (6)
Chemical	23	8	17	6	5	0
Computers and communications	16	5	12	2	3	1
Drugs and medical	18	9	12	7	4	3
Electrical and electronic	16	17	11	9	2	2
Mechanical	18	14	12	9	2	1
Others	30	18	21	11	7	1
Total	121	71	85	44	23	8
Max. no. of rejections	252	252	180	180	72	72
Percentage	0.48	0.28	0.47	0.24	0.32	0.11

[a]The number of rejections of one-sided 5% t-tests for individual coefficients is shown in each cell of the table for the early period (1975–1989) and for the late period (1990–1999) separately. Columns (1) and (2) show the number of rejections for all coefficients for country dummies (42 coefficients), columns (3) and (4) show the number of rejections for country dummy coefficients, dropping the rest of the world coefficients and also coefficients from rest of the world cited patent regressions (as a result, 30 coefficients), and columns (5) and (6) show the number of rejections for EU country dummy coefficients of EU cited patent regressions (hence, further reduced to 12 coefficients). The test results are based on the no-fixed-effect (no FE) estimator.

important to control for censoring in this comparison, as the results from the second period will be much more affected by censoring than the former period.

In columns (1) and (2) of Table 7, we see that there is a large decline in rejection rates over time. No home bias is rejected in 48% of cases in the early period, but only for 28% of cases in the later period (in the table, we report results at the 5% significance level). In columns (3) and (4), we repeat the exercise, but focus on OECD countries.[26] There is substantial home bias for the non-OECD countries, as noted above, so we wanted to check that the time series changes are not being driven by them alone. It is clear that the main patterns of results stand up. Although the absolute level of home bias is lower, the fall in the degree of home bias is dramatic. On average, the rejection rate falls from 47% to 24%. The final two columns look within the European countries (counting rejections only on European country dummy coefficients of European-country-cited patent regressions). The patterns are similar, with a large decline in home bias.

As we saw above, controlling for unobserved heterogeneity is important. In Table 8, we find that in most cases in both periods, the level of home bias is lower when we control for fixed effects (and censoring), but the reduction in home bias is more substantial in the early period than in the later period (the rejection rate falls from 30% to 22%). The reduction in home bias over time is less striking, because there is less evidence of home bias existing in the first place. Looking at the first two columns of Table 8, we see that

[26]In other words, we report the number of rejections for country dummy coefficients, dropping the rest of the world coefficients and also dropping coefficients from the rest of the world cited patent regressions.

Table 8. Number of rejections of no home bias using subsamples. The estimation method is the censored fixed-effect estimator.[a]

	All Countries		OECD Countries		EU Countries	
Technological Category	Early (1)	Late (2)	Early (3)	Late (4)	Early (5)	Late (6)
Chemical	13	7	9	4	1	0
Computers and communications	13	15	10	10	1	2
Drugs and medical	9	10	6	7	1	4
Electrical and electronic	12	6	9	5	0	1
Mechanical	15	8	12	5	4	1
Others	14	10	10	6	2	2
Total	76	56	56	37	9	10
Max. no. of rejections	252	252	180	180	72	72
Percentage	0.30	0.22	0.31	0.21	0.12	0.14

[a]The number of rejections of one-sided 5% t-tests for individual coefficients is shown in each cell of the table for the early period (1975–1989) and for the late period (1990–1999) separately. Columns (1) and (2) show the number of rejections for all coefficients for country dummies (42 coefficients), columns (3) and (4) show the number of rejections for country dummy coefficients, dropping the rest of the world coefficients and also coefficients from rest of the world cited patent regressions (as a result, 30 coefficients), and columns (5) and (6) show the number of rejections for EU country dummy coefficients of EU cited patent regressions (hence, further reduced to 12 coefficients). The test results are based on the censored fixed-effect (FE + C) estimator.

home bias declined in chemical, electrical and mechanical. By contrast, in computers and communication (ICT) and drugs, the modern sectors, we see little change (if anything an increase in the number of rejections). At first glance this might seem surprising as it is commonly assumed that ICT leads to delocalization.

Our aim in this paper is to identify "stylized facts" on home bias; what we identify is a reduced form of various structural influences that could slow down knowledge diffusion. What might these structural influences be? First, there are explicit information acquisition and communication costs that make it harder for inventors in country A to learn about inventions in country B because of telecommunication prices. The advent of e-mail, cellular phones, the Internet, liberalization of state telephone monopolies, and so forth has clearly reduced these explicit costs. In opposition to this, there are various agglomeration effects that will tend to make local interaction more important (at least in some sectors). When technologies are complex and/or at an early stage, then local communication to facilitate the transfer of tacit know-how may be particularly important.

ICT and pharmaceuticals are the two sectors where there has been the most discussion of "clustering" (e.g., ICT in Silicon Valley and biotechnology in Cambridge, Massachusetts).[27] These results are also shown in Figure 4, where the left-hand diagrams are of the early period and the right-hand diagrams are of the late period: the later period has far fewer circles (evidence for home bias) than the earlier period.

[27]For example, see Zucker, Darby, and Brewer (1998) on biotechnology.

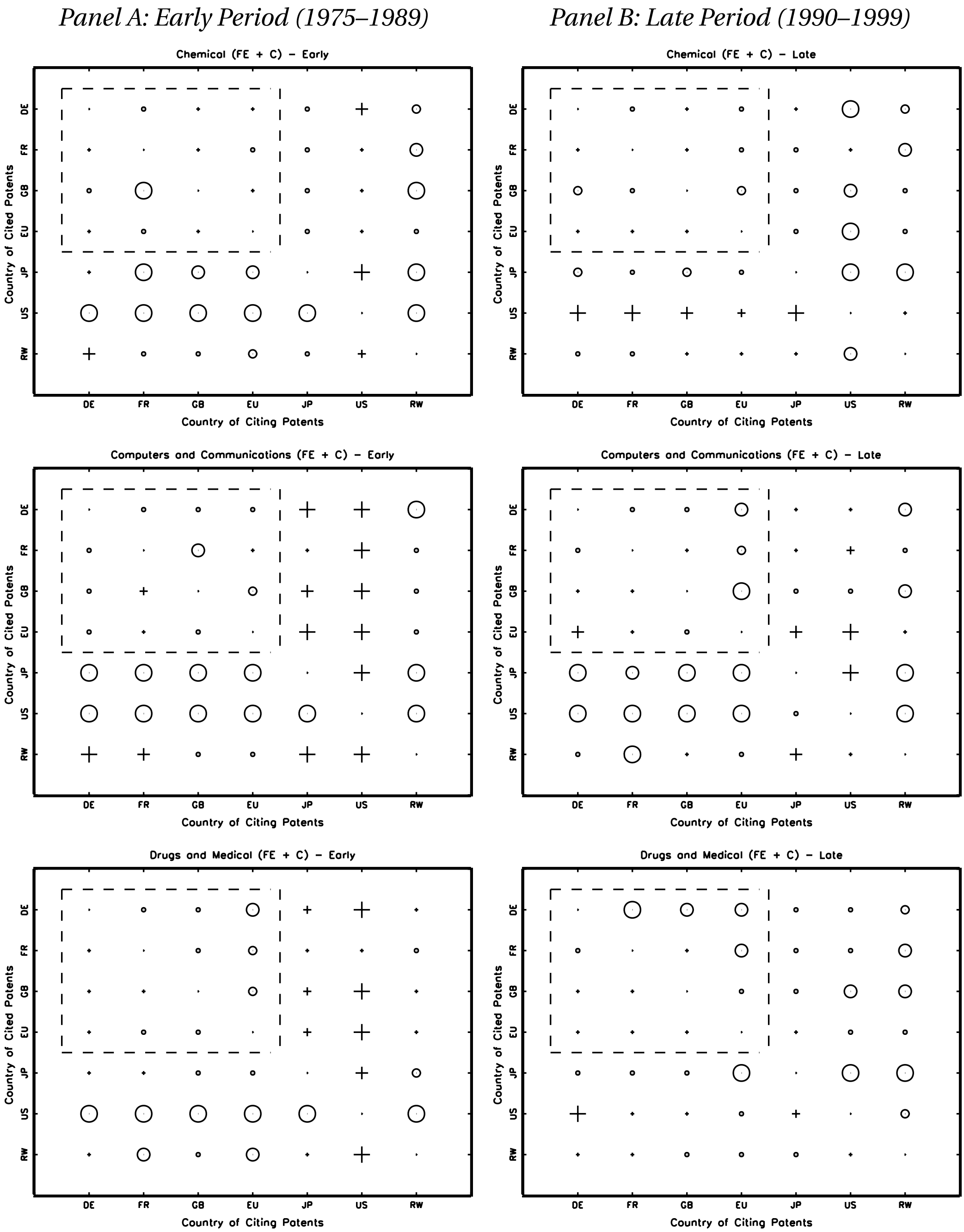

FIGURE 4. Early period versus late period. The left-hand diagrams are estimation results for the early period (1975–1989) and the right-hand diagrams are for the late period (1990–1999). Estimation results are from our preferred fixed-effects plus censoring specifications. The upper left quadrants with dashed boxes contain the cross-citations from the European countries.

Panel A: Early Period (1975–1989) *Panel B: Late Period (1990–1999)*

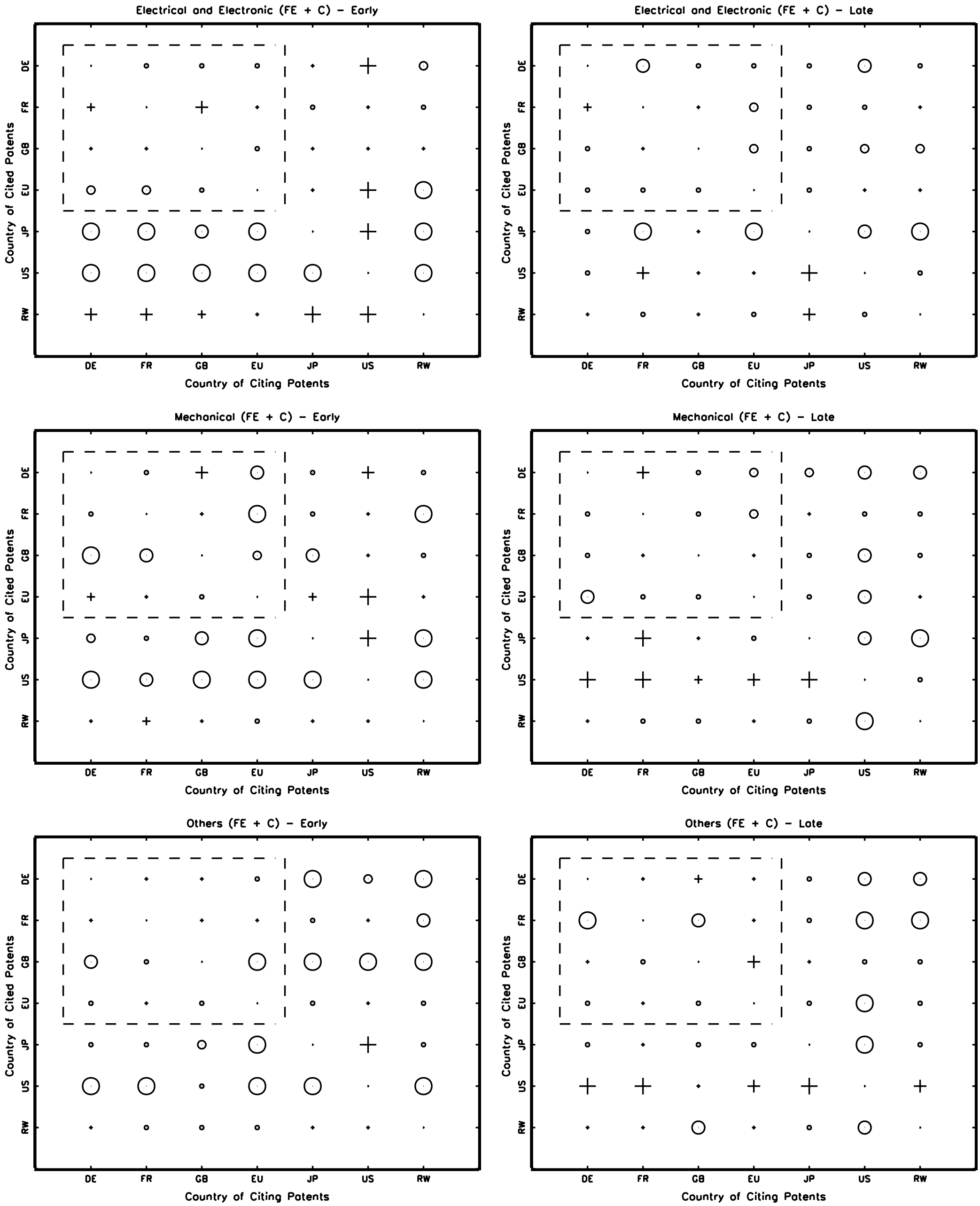

FIGURE 4. *Continued.*

It is plausible that in computers and communications, agglomeration effects may have become stronger over time (relative to other sectors) so as to offset the falling communication costs from which they would disproportionately benefit. It is hard to produce direct evidence to support this case, but there is some indirect evidence consistent with this idea. First, in terms of technological complexity, the ICT-producing industries have enjoyed very rapid technological change over our sample period. Quality-adjusted prices for computers were falling at something like 15% per annum until the early 1990s, after which the price declines accelerated to around 30%. This appeared to be partially driven by a speedup in the technological cycle of the semiconductor industry.[28] Rapid technological change is thought to make face-to-face communication and geographically based knowledge spillovers more important.

Second, the geographical clustering of the ICT industry appears to have strengthened rather than weakened over time. For example, despite high wages and land prices in Silicon Valley, leading software companies such as Apple, Oracle, and Google have not chosen to decamp en masse, but seem to benefit from proximity to other ICT-oriented firms.

Third, several recent papers have pointed to the rise of "superstar cities," which have highly skilled workers and high-tech industries, and workers increasingly collocated in the same localities (for example, Gyourko, Mayer, and Sinai (2006)). These cities (such as San Francisco and New York) have an increasing concentration of graduate workers and high-tech jobs, and have lead to increasing spatial inequality. There are various hypotheses that could explain this, but one leading explanation is that high-tech sectors, such as ICT, are increasingly clustered in certain areas and this generates demand for high skilled workers (for example, Moretti (2010)).

While we cannot be certain that these are the reasons that there we do not see a fall in home bias in some industries, the most likely rationalization seems to be an increased importance of agglomeration has offset the fall in explicit communication costs along the lines discussed.

One of the things that Figure 4 shows quite clearly is that in the early period, the United States had a substantial advantage in terms of "absorptive capacity": U.S. firms were quick to learn from and cite European patents, while European firms were slower to cite U.S. firms and each other. In the later period, European firms now cite U.S. patents more quickly, and the United States has lost this advantage. We see this by comparing the US row for U.S. cited patents in the early period (on the left-hand side) with the later period (the US row on the right-hand side). The evidence of home bias (large circles) has been replaced by evidence of no home bias (a dot) or, in some cases, even evidence of more speedy knowledge flows (indicated by pluses). This is true in all industries except computers, where the United States seems to have maintained its advantage in terms of absorptive capacity.

The obvious conclusion is that home bias has fallen, and it has fallen in those sectors where one would a priori expect it to fall. This seems to be new concrete quantitative evidence for an aspect of globalization that is much discussed—the increasing propensity of knowledge to slip over geographic boundaries. Our conclusion is consistent with Keller (2002), who showed that geographic localization declined over time between countries, using a model in which productivity depends on domestic and foreign R&D, and the effectiveness of foreign R&D is negatively related to the distance

[28]Some authors go so far as to say that the productivity acceleration in the United States was in large part due to accelerated technical progress in ICT (Jorgenson (2001) and van Ark, O'Mahony, and Timmer (2008)).

from the foreign economy. The conclusion also is consistent with recent evidence from Kim, Morse, and Zingales (2006) that the spillover benefits that academics obtain from their colleagues within the same university are less important now than they were two decades ago.

4.3 *Using longer lags of citations*

Our baseline results use only the first two citations to measure home bias. Why not use the third, fourth, or fifth citation? Our main reason is because we believe that the theoretically relevant information is contained in the first few citations. This is before the patent has become more general public knowledge; it is when information is the most tacit. After the patent has been published and cited, it becomes codified and there is less reason to believe that geography should matter. In addition, we have argued for a smaller number of citations on the grounds of theory (the first few cites are likely to be where home bias is greatest due to tacitness of knowledge) and parsimony (we need at least two observations to difference out the fixed effect, so the first two citations comprise the minimum number).

Nevertheless, to tackle this issue directly, we also checked the robustness of our results to including the third and fourth cites. The conditional likelihood estimator developed in Section 2 can easily be extended for more than two citations. Suppose that $J = 3$, that is, that there are three potential citing patents. Then it is straightforward to show that the probability that the observed second citation is second, conditional on the durations of the first and second citations, has the logit form as in equation (1), independent of unobserved heterogeneity. Thus, this implies that one can obtain another censored fixed-effect estimator in exactly the same way as in equation (2) by replacing the subscripts 1 and 2 with subscript 2 and 3, respectively.[29]

Our qualitative findings did not change.[30] For example, in Table 6, for the 5% level, the number of rejections falls from 150 (no FE) to 77 (FE + C) as we control for unobserved heterogeneity of citing patents and the censoring problem. When we use the second and third citations for the same level, the number of rejections changes from 130 (no FE) to 75 (FE + C); when we use the third and fourth citations, the number falls from 121 (no FE) to 48 (FE + C). The larger decline with the third and fourth citations is consistent with our conjecture that geography is less important as the patent becomes more general public knowledge.

4.4 *The importance of controlling for base of citing patents*

Our specifications include a control for the number of citing patents (the base) by country and technology class. Although this might seem natural, in our econometric specification, we may not need to control for it because the partial likelihood estimator is

[29] Similarly, if $J = 4$, one can show that the probability that the observed third citation is third, conditional on the durations of the first, second, and third citations has the logit form again, independent of unobserved heterogeneity. Then one can obtain yet another censored fixed estimator exactly the same way as in equation (2) by replacing the subscripts 1 and 2 with subscript 3 and 4, respectively.

[30] See Tables A.7 and A.8 in Appendix B.

based on the partial likelihood of the identity (including origin country) of the first citation, given the time of this first citation and the identities of the first two citations. However, our alternative interpretation is that legal and organizational changes to the patent system led to large increases in patenting in some technology class–country pairs that led to falls in average patent quality. Indeed, the coefficient on base is robustly negative, consistent with this interpretation.

We dropped base and reestimated all our regressions, which led in the main to qualitatively similar results. In particular, we found evidence of home bias weakening both with fixed effects and over time. We did find, however, that other countries appeared to cite U.S. inventors more quickly than inventors in their own countries when we failed to control for base (e.g., more crosses in the row marked US in Table 3). We suspect that this is because many countries took advantage of looser rules on U.S. patenting to expand their portfolios and also cited more out of fear of litigation. Conditioning on the total number of cited patents controls for this bias and restores the result that (in general) other countries cite themselves more quickly than they do U.S. inventors.

4.5 *Further robustness tests*

Could there be other reasons why the apparent decline in home bias is spurious? First, a concern may be that the number of rejections of home bias has fallen because the number of observations is lower in the late period. But Table 2 shows that, if anything, the number of patents is slightly higher in the later period (1.107 million vs. 1.031 million), so this cannot be the reason. Second, could it be that the differential quality of patents has caused this to occur? For example, a lot of the decline in Figure 4 is because European firms have become relatively faster at citing U.S. patents, and Japanese firms have become relatively slower. Our technique of using multiple cites to difference out the fixed effect means that we have controlled for cited patent quality. Consequently, differential quality cannot be the reason for the patterns we observe in Table 8 (but it might be the reason for the patterns observed in Table 7, which does not control for fixed effects). Third, we also tried using different cutoff years and found that this led to similar results. For example, we obtain qualitatively similar results using 1985 as a cutoff year with the censored fixed-effect estimator: in the chemical category, the number of rejections using all countries decreased from 13 in the pre-1985 period to 8 in the post-1985 period; in mechanical engineering, the fall was from 12 to 5.[31] See Table A.6 in Appendix B for details.

A final concern is that our results might be driven by self-citations. Of course, the positive coefficient on self-citations may reflect some degree of localized knowledge spillovers and so is of interest in its own right. Nevertheless, self-citations could reflect a bias simply to overcite oneself, so we dropped self-citations and reestimated all models. This made little difference to the preferred results with fixed effects and censoring. For example, the number of rejections at the 5% level in Table 6 fell merely from 77 to 76. Dropping self-citations *did* make more of a difference for the no-fixed-effects models,

[31]As before, the modern sectors have seen an increase in the number of rejections from 7 to 9 in pharmaceuticals and from 9 to 11 in electrical and electronic.

however, with the number of rejections falling from 150 in column (2) to 118 when we drop self-citations. We conclude that another benefit of our methodology is that it helps correct for biases induced from self-citations. See Table A.9 in Appendix B for details.

5. Conclusions

Patent citations have become an important source of information about the ways in which knowledge flows between firms and countries. But knowledge can spread more or less quickly due to the unobservable characteristics of patents, which may be poorly captured by observable characteristics. In this paper, we propose an econometric technique for dealing with fixed effects in duration models that exploits the existence of multiple citations on the same patent and implements this estimator on a data base of over two million citations between 1975 and 1999. We have focussed on the speed of knowledge flows between countries, which is a key feature of models of growth and international trade. Many papers have argued that there is substantial home bias in the way that knowledge is transmitted, in the sense that being geographically close makes knowledge transfers easier, and this has become accepted wisdom in government support for clusters and other forms of technology policy.

We find that controlling for unobserved heterogeneity makes a large quantitative and qualitative difference to estimates of home bias in innovative activity. First, the evidence for home bias is much weaker once we control for fixed effects (and censoring). The non-fixed-effects models (which are standard in the literature) suggest home bias in a majority of cases, whereas our preferred models indicate home bias in only a minority of cases. Second, and perhaps most provocatively, we find evidence that home bias has declined over time, being much stronger in the pre-1990 period than the post-1990 period. We interpret this as suggesting that information flows more easily across national boundaries as the cost of international communication and travel has fallen. Furthermore, there is heterogeneity in the fall in home bias: it has not occurred in the more high-tech sectors of ICT and pharmaceuticals, precisely those areas where clusters and agglomeration are believed to be important. This suggests that international boundaries may be less important, but that in many sectors distance is far from dead.

Appendix A: Econometric appendix

A.1 *Likelihood function with censoring*

The censoring time C_i for patent i is defined as the number of days from the date of patent i being granted to the common censoring date. We assume that the censoring time C_i is independent (Y_{ij}^*, X_{ij}, U_i) and identically distributed with an unknown probability distribution. Furthermore, we assume that the support of C_i is the whole real line. Under this censoring mechanism, our data consist of $\{(Y_{ij}, \Delta_{ij}, X_{ij}, C_i) : i = 1, \dots, n, j = 1, \dots, J\}$, where $Y_{ij} = \min(Y_{ij}^*, C_i)$ and $\Delta_{ij} = 1(Y_{ij}^* < C_i)$. Here, $1(\cdot)$ is the usual indicator function. Thus, we observe uncensored citation durations only when $\Delta_{ij} = 1$, that is, citation durations are less than the censoring time.

In this paper, we propose a modified version of the conditional likelihood estimator to correct for the selection bias. Specifically, the proposed estimator of β, say $\hat{\beta}$, maximizes the following weighted conditional log-likelihood function with $J = 2$:

$$
\begin{aligned}
L(b) = n^{-1} \sum_{i=1}^{n} & \frac{\Delta_{i1}\Delta_{i2}}{G_n(\max[Y_{i1}, Y_{i2}])} \\
& \times \left\{ \left[1(Y_{i1} \leq Y_{i2}) \ln\left(\frac{\exp(X_{i1}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)} \right) \right] \right. \\
& \left. + 1(Y_{i1} \geq Y_{i2}) \ln\left[\frac{\exp(X_{i2}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)} \right] \right\},
\end{aligned}
\tag{A.1}
$$

where $G_n(\cdot)$ is an estimator of the survivor function $G(\cdot)$ of the censoring time C_i; in particular, $G_n(c) = n^{-1} \sum_{i=1}^{n} 1(C_i > c)$. Our econometric framework is based on a continuous-time duration model, which is suitable for our application since we have citation durations measured in days. However, it is possible to have ties and they are included in both contributed terms in (A.1). Observe that the selection bias is corrected for by multiplying weights $G_n(\max\{Y_{i1}, Y_{i2}\})^{-1}$ in equation (A.1). The reason why $G_n(\max\{Y_{i1}, Y_{i2}\})^{-1}$ represents proper weights is that

$$
E\left[\frac{\Delta_{i1}\Delta_{i2}}{G(\max\{Y_{i1}, Y_{i2}\})} \middle| Y_{i1}^*, Y_{i2}^*, X_{i1}, X_{i2} \right] = 1. \tag{A.2}
$$

In other words, (A.1) converges in probability uniformly over b to a limiting function to which an infeasible log-likelihood function would converge under no censoring. In maximizing (A.1), we trim away 0.5% of observations with the smallest values of $G_n(\max[Y_{i1}, Y_{i2}])$ to mitigate the leverage of outliers.

A.2 *Asymptotic distribution of the censored fixed-effect estimator*

This section of the appendix describes regularity conditions under which the censored fixed-effect estimator is consistent and asymptotically normal. Also, it gives the form of asymptotic variance of the censored fixed-effect estimator.

Assumption A.1. (i) *β is an interior point of a compact subset of $\mathbf{R}^d$ for some finite d.* (ii) *The data $\{(Y_{i1}, Y_{i2}, X_{i1}, X_{i2}, \Delta_{i1}, \Delta_{i2}, C_i) : i = 1, \dots, n\}$ are independent and identically distributed.* (iii) *Y_{i1}^* and Y_{i2}^* are independent of each other conditional on (X_{i1}, X_{i2}, U_i).* (iv) *$\lambda_i(\cdot)$ is strictly positive.* (v) *$E[\|X_{i1} - X_{i2}\|^2] < \infty$ and $E[(X_{i1} - X_{i2})(X_{i1} - X_{i2})']$ is nonsingular.* (vi) *The censoring variable C_i is random with an unknown continuous probability distribution.* (vii) *C_i is independent of $(Y_{i1}^*, Y_{i2}^*, X_{i1}, X_{i2}, U_i)$.* (viii) *The survivor function of C_i, $G(c) \equiv \Pr(C_i > c)$, is positive for every $c \in \mathbf{R}$ and, furthermore, it is bounded away from zero.*

These assumptions are not unrestrictive, but in our application, they might be viewed as plausible. Recall that citation durations are obtained by looking at all recorded citations as of December 31, 1999. Hence, the censoring variable is defined as the difference between this particular end date and the date when a patent was granted. It is reasonable that the censoring time C_i is independent of potential citation durations Y_{ij}^*, the attributes of the citing patent X_{ij}, and the heterogeneity term U_i, because the dates of patents being granted may have little to do with underlying patent-citing processes.[32] Also, the full support condition (viii) on the censoring time is not so restrictive in our application given that we follow patent citations over a long period and we focus mainly on the first two citations. The assumption that $G(\cdot)$ is bounded away from zero is useful to ensure that our estimator behaves regularly. For example, see Khan and Tamer (2010) for general issues regarding inverse weight estimation.

Let

$$\begin{aligned} H_i(b) = {} & 1(Y_{i1} \leq Y_{i2})[X_{i1} - X_{i2}] \frac{\exp(X_{i2}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)} \\ & + 1(Y_{i1} \geq Y_{i2})[X_{i2} - X_{i1}] \frac{\exp(X_{i1}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)}. \end{aligned} \tag{A.3}$$

Define

$$\Omega = \Gamma^{-1}\left\{\operatorname{Var}\left[\frac{\Delta_1 \Delta_2}{G(\max\{Y_1, Y_2\})} H(\beta)\right] - \operatorname{Var}[\rho(C)]\right\}\Gamma^{-1},$$

where

$$\Gamma = E\left[-\frac{\partial^2 L(\beta)}{\partial b\, \partial b'}\right] \quad \text{and} \quad \rho(c) = E\left[\frac{\Delta_1 \Delta_2 H(\beta)}{G^2(\max\{Y_1, Y_2\})} 1(c > \max\{Y_1, Y_2\})\right].$$

Then the following theorem gives the asymptotic normality of the censored fixed-effect estimator.

Theorem A.1. *Let Assumption A.1 hold. Assume that Ω exists and is finite. Then as $n \to \infty$,*

$$\sqrt{n}(\hat{\beta} - \beta) \to_d \mathbf{N}(0, \Omega). \tag{A.4}$$

The proof of Theorem A.1 is omitted; it can be proved as in the proof of Theorem 1 of Lee (2008). The asymptotic variance Ω can be consistently estimated by

$$\hat{\Omega} = \hat{\Gamma}^{-1}\left[n^{-1}\sum_{i=1}^{n}(\hat{\Phi}_i - \hat{\rho}_i)(\hat{\Phi}_i - \hat{\rho}_i)'\right]\hat{\Gamma}^{-1},$$

[32]What we need is that the application and grant dates are independent of quality. However, the restriction that the application date is independent of quality can be violated if there is a cohort effect on cited patents such as technology waves. Another problematic case would be if the time lag between the application date and the grant date is systematically correlated with the quality of the patent. Then this would induce the dependence between the grant date and quality, even when the application date is exogenous.

where $G_{ni} = G_n(\max\{Y_{i1}, Y_{i2}\})$,

$$\hat{\Gamma} = n^{-1} \sum_{i=1}^{n} \frac{\Delta_{i1}\Delta_{i2}}{G_{ni}} [X_{i1} - X_{i2}][X_{i1} - X_{i2}]' \frac{\exp(X'_{i1}\hat{\beta} + X'_{i2}\hat{\beta})}{[\exp(X'_{i1}\hat{\beta}) + \exp(X'_{i2}\hat{\beta})]^2},$$

$$\hat{\Phi}_i = \frac{\Delta_{i1}\Delta_{i2}}{G_{ni}} H_i(\hat{\beta}),$$

and

$$\hat{\rho}_i = n^{-1} \sum_{k=1}^{n} \left[\frac{\Delta_{1k}\Delta_{2k} H_k(\hat{\beta})}{G_{nk}^2} 1(C_i > \max\{Y_{1k}, Y_{2k}\}) \right].$$

A.3 *Related econometric models in the literature*

A.3.1 *Jaffe and Trajtenberg* In Jaffe and Trajtenberg (1999), the likelihood that a particular patent K (citing patent) granted in year T will cite some patent k granted in year t (cited patent) has the form

$$\alpha(k, K) \cdot \exp[-\beta_1(k, K) \cdot (T - t)] \cdot \{1 - \exp[-\beta_2 \cdot (T - t)]\},$$

where α is a shift parameter that depends on the attributes of patents k and K, β_1 is an obsolescence parameter that also depends on the characteristics of patents k and K, and β_2 is a diffusion parameter. The first exponential process, $\exp[-\beta_1(T - t)]$, describes how knowledge becomes obsolete and the second exponential process, $1 - \exp[-\beta_2(T - t)]$, models how knowledge diffuses. Since we focus on the first few citations, the aspect of knowledge obsolescence is far less important in our empirical work than in Jaffe and Trajtenberg (1999). Roughly speaking, a natural form of specializing the citation frequency of Jaffe and Trajtenberg (1999) to our setup would be

$$P_{\mathrm{JT}} := \alpha(k, K) \cdot \{1 - \exp[-\beta_2 \cdot (T - t)]\}. \tag{A.5}$$

Note that our mixed proportional hazards model specification gives the citation frequency

$$P_{\mathrm{MPH}} := 1 - \exp[-\Lambda_i(T - t) \exp(x'_{ij}\beta + u_i)], \tag{A.6}$$

where $\Lambda_i(u) := \int_0^u \lambda_i(s)\,ds$ is the integrated baseline function. The Jaffe–Trajtenberg-style model in (A.5) assumes proportionality in terms of the citation frequency P_{JT}; however, our mixed proportional hazards model in (A.6) takes the proportionality in terms of the hazard function. In general, these two models are nonnested; however, if we assume that $\alpha(k, K) \equiv 1$ but β_2 may depend on x_{ij} and u_i as in (A.6), then (A.5) is a special case of (A.6) with $\Lambda_i(u) = u$ (no duration dependence in the baseline hazards).

As we mentioned in the main text, we control for unobserved heterogeneity in a way that Jaffe and Trajtenberg (1999) do not. Still, we see our approach as a complement rather than a substitute for Jaffe and Trajtenberg (1999) since, strictly speaking, both models are nonnested.

A.3.2 *Thompson* Thompson (2006) reported estimates from conditional logits with fixed effects for each cited patent. Although our estimates are also from a conditional logit with fixed effects for cited patents, two methods are quite distinct.

First of all, we need to worry about the problem of censoring, since our framework is based on a duration model; however, Thompson (2006) was free from censoring problems, since he considered observed pairs of cited–citing patents for the conditional logit. Furthermore, our conditional logit estimates use an indicator whether citation from an inventor residing in the same country has a shorter duration as the dependent variable and country dummies as important explanatory variables; whereas Thompson's logit estimates use an indicator whether both inventors of the cited–citing patents reside in the same country as the dependent variable and an indicator variable whether the citation is added by the inventor as the main explanatory variable.

Thompson's (2006, Table 3) estimation results suggest that the localization of international knowledge spillovers has not declined over time. However, two different estimation results are associated with different samples of patent citation data. Thompson's sample starts from all patents granted during the first week of January 2003 and having an institutional assignee, and then pairs of cited–citing patents are constructed by all patents cited in this particular cohort of citing patents. Our sample consists of potentially cited patents between 1975 and 1999, and its corresponding first few citing patents. In short, Thompson's data extract is based on citing patents, whereas our data extract is based on cited patents.

Appendix B: Additional data description and results

In this appendix we include several tables that show additional results.

Table A.1 shows a tabulation of the country of the first patent citing each of the cited patents in our data. The diagonal elements show that there is substantial home bias in the raw data. A problem we face in evaluating the time taken until the first patent is that not all patents have been cited. Estimating on only those patents where we observe two citations would lead to potential selection bias. Table A.2 shows the number of patents that are censored, by industry; Table A.3 splits this down into the early and late period,

Table A.1. Raw data: home bias in first citation.[a]

	Citing						
Cited	DE	FR	GB	EU	JP	US	RW
DE	**30.58**	2.96	2.90	5.05	14.67	38.82	5.03
FR	8.56	**19.00**	3.37	5.24	13.29	45.18	5.36
GB	8.36	3.10	**16.57**	4.57	12.76	49.46	5.17
EU	8.98	3.41	2.81	**21.21**	13.15	44.18	6.25
JP	5.47	1.68	1.59	2.74	**50.90**	33.05	4.56
US	5.03	2.04	2.29	3.12	10.86	**71.83**	4.82
RW	7.17	2.64	2.47	4.63	12.49	49.00	**21.61**

[a]Data consist of all patents that were granted between 1975 and 1999 (the cited patent) and the first patent to cite it (the citing patent). An element $\{i, j\}$ in the table shows the proportion of patents granted to an inventor located in row country i that are first cited by an inventor in a column country j. For example, element $\{1, 2\}$ indicates that 2.96% of patents from German inventors were first cited by an inventor in France.

TABLE A.2. Censoring: many patents have not (yet) been cited.[a]

Observed	Chemicals	Computer	Drugs	Electrical	Mechanical	Other	Total
2 cites	241,799	152,557	90,718	220,584	250,258	217,545	1,173,461
	(56.35)	(61.75)	(49.48)	(60.14)	(53.54)	(48.89)	(54.87)
1 cite	65,969	29,483	25,348	52,985	78,875	774,91	330,151
	(15.37)	(11.93)	(13.83)	(14.45)	(16.87)	(17.41)	(15.44)
No cites	121,364	65,009	67,274	93,192	138,318	149,939	635,096
	(28.28)	(26.31)	(36.69)	(25.41)	(29.59)	(33.70)	(29.70)

[a]Each row indicates the number of observations that had at least two cites, one cite, or no cites. The numbers in parentheses indicate the proportion of observations by industry that had different numbers of cites. For example, our data set contains 254,301 cites to patents in the chemical technology sector that had at least two cites.

TABLE A.3. Censoring by early and late time period.[a]

Observed	Chemicals	Computer	Drugs	Electrical	Mechanical	Other	Total
			1975–1989				
2 cites	166,715	69,718	48,007	133,242	169,426	151,371	738,479
	(72.08)	(86.36)	(73.32)	(78.28)	(68.07)	(64.65)	(71.64)
1 cite	31,987	6324	8268	20,397	39,596	39,510	146,082
	(13.83)	(7.83)	(12.63)	(11.98)	(15.91)	(16.87)	(14.17)
No cites	32,596	4685	9204	16,583	39,868	43,260	146,196
	(14.09)	(5.80)	(14.06)	(9.74)	(16.02)	(18.48)	(14.18)
			1990–1999				
2 cites	75,084	82,839	42,711	87,342	80,832	66,174	434,982
	(37.95)	(49.81)	(36.24)	(44.44)	(36.98)	(31.39)	(39.26)
1 cite	33,982	23,159	17,080	32,588	39,279	37,981	184,069
	(17.18)	(13.92)	(14.49)	(16.58)	(17.97)	(18.01)	(16.61)
No cites	88,768	60,324	58,070	76,609	98,450	106,679	488,900
	(44.87)	(36.27)	(49.27)	(38.98)	(45.04)	(50.60)	(44.13)

[a]This table is the same as Table A.2 except we now split into early and later years.

TABLE A.4. Censoring by cited country.[a]

	Cited Country							
Observed	DE	FR	GB	EU	JP	US	RW	Total
2 cites	91,587	33,852	36,684	49,356	229,321	668,492	64,169	1,173,461
	(53.94)	(52.07)	(56.39)	(49.73)	(58.70)	(55.71)	(43.06)	(54.87)
1 cite	28,724	11,170	10,337	16,955	56,077	180,929	25,959	330,151
	(16.92)	(17.18)	(15.89)	(17.08)	(14.36)	(15.08)	(17.42)	(15.44)
No cites	49,496	19,992	18,030	32,938	105,246	350,511	58,883	635,096
	(29.15)	(30.75)	(27.72)	(33.19)	(26.94)	(29.21)	(39.52)	(29.70)

[a]This table is the same as Table A.2 except we now split into countries.

clearly showing that the censoring problem is much more significant in the later period; Table A.4 shows this by cited country. This motivates our use of estimators that explicitly allow for censoring.

TABLE A.5. Number of rejections of no home bias using subsamples. The estimation method is the no-fixed-effect estimator. The cutoff year is 1985.[a]

	All Countries		OECD Countries		EU Countries	
Technological Category	Early (1)	Late (2)	Early (3)	Late (4)	Early (5)	Late (6)
Chemical	24	20	18	15	6	5
Computers and communications	14	9	11	5	2	1
Drugs and medical	15	14	8	8	2	4
Electrical and electronic	16	18	10	10	1	1
Mechanical	19	23	13	14	3	4
Others	29	24	21	14	5	3
Total	117	108	81	66	19	18
Max. no. of rejections	252	252	180	180	72	72
Percentage	0.46	0.43	0.45	0.37	0.26	0.25

[a]The number of rejections of one-sided 5% t-tests for individual coefficients is shown in each cell of the table for the early period (1975–1984) and for the late period (1985–1999) separately. Note that the tables in the main text use 1990 as the cutoff year. Columns (1) and (2) show the number of rejections for all coefficients for country dummies (42 coefficients), columns (3) and (4) show the number of rejections for country dummy coefficients, dropping the rest of the world coefficients and also coefficients from rest of the world cited patent regressions (as a result, 30 coefficients), and columns (5) and (6) show the number of rejections for EU country dummy coefficients of EU cited patent regressions (hence, further reduced to 12 coefficients). The test results are based on the no-fixed-effect (no FE) estimator.

TABLE A.6. Number of rejections of no home bias using subsamples. The estimation method is the censored fixed-effect estimator. The cutoff year is 1985.[a]

	All Countries		OECD Countries		EU Countries	
Technological Category	Early (1)	Late (2)	Early (3)	Late (4)	Early (5)	Late (6)
Chemical	13	8	8	4	1	0
Computers and communications	12	13	9	10	1	1
Drugs and medical	7	9	5	6	0	2
Electrical and electronic	9	11	7	7	0	1
Mechanical	12	5	9	2	3	0
Others	13	12	9	9	2	3
Total	66	58	47	38	7	7
Max. no. of rejections	252	252	180	180	72	72
Percentage	0.26	0.23	0.26	0.21	0.10	0.10

[a]The number of rejections of one-sided 5% t-tests for individual coefficients is shown in each cell of the table for the early period (1975–1984) and for the late period (1985–1999) separately. Note that the tables in the main text use 1990 as the cutoff year. Columns (1) and (2) show the number of rejections for all coefficients for country dummies (42 coefficients), columns (3) and (4) show the number of rejections for country dummy coefficients, dropping the rest of the world coefficients and also coefficients from rest of the world cited patent regressions (as a result, 30 coefficients), and columns (5) and (6) show the number of rejections for EU country dummy coefficients of EU cited patent regressions (hence, further reduced to 12 coefficients). The test results are based on the censored fixed-effect (FE + C) estimator.

TABLE A.7. Number of rejections of no home bias using the entire sample with second and third citation spells.[a]

Technological Category	No FE			FE			FE + C		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
Chemical	30	26	19	13	12	5	12	8	3
Computers and communications	26	20	15	15	14	8	18	15	8
Drugs and medical	25	19	7	16	12	9	15	13	11
Electrical and electronic	20	16	15	17	12	9	21	16	10
Mechanical	26	22	19	22	17	11	17	14	9
Others	29	27	21	12	8	5	12	9	7
Total	156	130	96	95	75	47	95	75	48

[a]This table is the equivalent of Table 6 in the main text except we use estimates based on the second and third citations (instead of the first and second citations).

TABLE A.8. Number of rejections of no home bias using the entire sample with third and fourth citation spells.[a]

Technological Category	No FE			FE			FE + C		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
Chemical	28	22	17	13	11	6	11	8	4
Computers and communications	20	17	11	11	10	9	13	11	8
Drugs and medical	22	14	7	12	9	4	12	8	6
Electrical and electronic	24	20	16	14	11	8	17	11	8
Mechanical	23	20	16	8	6	3	9	5	3
Others	29	28	20	12	7	4	7	5	2
Total	146	121	87	70	54	34	69	48	31

[a]This table is the equivalent of Table 6 in the main text except we use estimates based on the third and fourth citations (instead of the first and second citations).

In investigating the change in home bias over time, we chose 1990 as a cutoff year because this approximately balances the number of citations in early and later years. In Tables A.5 and A.6, we show the robustness of the results to using the middle year of our sample period, 1985. As also discussed in the main test, we focus on the first two citations for a patent. We can easily extend our method also using the third citation and quasidifference between the second and third citation, and we show the results from doing this in Table A.7. Similarly, we can use up to the fourth citation (see Table A.8). Table A.9 provides estimation results after dropping all self-citations. Our results are robust to using these alternative citations.

Table A.9. Number of rejections of no home bias using the entire sample (without self-citations).[a]

Technological Category	No FE			FE			FE + C		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
Chemical	27	25	21	16	12	9	14	14	9
Computers and communications	16	13	6	14	13	12	17	15	12
Drugs and medical	19	18	11	14	10	5	13	10	6
Electrical and electronic	20	16	15	13	11	11	13	13	10
Mechanical	20	19	10	14	11	7	15	13	8
Others	30	27	23	17	11	8	14	11	10
Total	132	118	86	88	68	52	86	76	55

[a]This table is the equivalent of Table 6 in the main text except we drop self-citations and reestimate all models.

References

van Ark, B., M. O'Mahony, and M. Timmer (2008), "The productivity gap between Europe and the United States: Trends and causes." *Journal of Economic Perspectives*, 22 (1), 25–44. [235]

Arora, A., R. Landau, and N. Rosenberg (1999), "Dynamics of comparative advantage in chemical industry." In *Sources of Industrial Leadership: Studies in Seven Industries* (D. Mowery and R. Nelson, eds.), Chapter 6, Cambridge University Press, Cambridge, U.K. [227]

Belenzon, S. and J. Van Reenen (2007), "How international citation propensities have changed over time: Has America become better at absorbing knowledge?" Mimeo, Centre for Economic Performance, London School of Economics. [220]

Bernstein, J. and P. Mohnen (1998), "International R&D spillovers between US and Japanese R&D intensive sectors." *Journal of International Economics*, 44, 315–338. [215]

Bloom, N. and J. Van Reenen (2002), "Real options, patents, productivity and market value: Evidence from a panel of British firms." *Economic Journal*, 112 (478), C97–C116. [221]

Branstetter, L. (2001), "Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the U.S. and Japan." *Journal of International Economics*, 53, 53–79. [215]

Branstetter, L. and M. Sakakibara (2002), "When do research consortia work well and why? Evidence from Japanese panel data." *American Economic Review*, 92 (1), 143–159. [215]

Cabellero, 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*, 1–64. [217]

Cairncross, F. (1997), *The Death of Distance: How the Communications Revolution Will Change Our Lives*. Harvard Business School Press, Harvard. [212]

Chamberlain, G. (1980), "Analysis of covariance with qualitative data." *Review of Economic Studies*, 47, 225–238. [219]

Chamberlain, G. (1985), "Heterogeneity, omitted variable bias, and duration dependence." In *Longitudinal Analysis of Labor Market Data* (J. J. Heckman and B. Singer, eds.), Cambridge University Press, Cambridge. [216, 218, 219, 220]

Coe, D. and E. Helpman (1995), "International R&D Spillovers." *European Economic Review*, 39, 859–887. [215]

Coyle, D. (1997), *Weightless World*. Princeton University Press, Princeton. [212]

Disdier, A.-C. and K. Head (2008), "The puzzling persistence of the distance effect on bilateral trade." *Review of Economics and Statistics*, 90 (1), 37–418. [212]

Evans, C. and J. Harrigan (2005), "Distance, time and specialization: Lean retailing in general equilibrium." *American Economic Review*, 95, 292–313. [212]

Friedman, T. (2005), *The World Is Flat*. Farrar, Strauss and Giroux, New York. [212]

Griffith, R., R. Harrison, and J. Van Reenen (2006), "How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing." *American Economic Review*, 96 (5), 1859–1875. [215]

Griffith, R., S. Redding, and J. Van Reenen (2004), "Mapping the two faces of R&D." *Review of Economics and Statistics*, 86 (4), 883–895. [215]

Griliches, Z. (1992), "The search for R&D spillovers." *Scandinavian Journal of Economics*, 94, S29–S47. [215]

Gyourko, J., C. Mayer, and T. Sinai (2006), "Superstar cities." Working Paper 12355, NBER. [235]

Hall, B., A. Jaffe, and M. Trajtenberg (2001), "The NBER patent citation data file: Lessons, insights and methodological tools." Working Paper 8498, NBER, see also http://www.nber.org/patents/. [220, 221]

Hall, B., A. Jaffe, and M. Trajtenberg (2005), "Market value and patent citations: A first look." *RAND Journal of Economics*, 36, 16–38. [220, 227]

Head, K., T. Mayer, and J. Ries (2007), "How remote is the offshoring threat?" Mimeo, University of Toronto. [216]

Henderson, R., A. Jaffe, and M. Trajtenberg (2005), "Patent citations and the geography of spillovers: A reassessment—comment." *American Economic Review*, 95 (1), 461–464. [215]

Horowitz, J. L. and S. Lee (2004), "Semiparametric estimation of a panel data proportional hazards model with fixed effects." *Journal of Econometrics*, 119, 155–198. [218]

Jaffe, A. B. (1986), "Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value." *American Economic Review*, 76 (5), 984–1001. [220]

Jaffe, A. and J. Lerner (2004), *Innovation and Its Discontents: How Our Broken Patent System Is Endangering Innovation and Progress, and What to Do About It*. Princeton University Press, Princeton. [222]

Jaffe, A. and M. Trajtenberg (1999), "International knowledge flows: Evidence from patent citations." *Economics of Innovation and New Technology*, 8 (1–2), 105–136. [215, 220, 227, 241]

Jaffe, A. and M. Trajtenberg (2002), *Patents, Citations and Innovation: A Window on the Knowledge Economy*. MIT Press, Cambridge. [220]

Jaffe, A., M. Trajtenberg, and R. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations." *Quarterly Journal of Economics*, 108 (3), 577–598. [215, 220]

Jorgenson, D. (2001), "Information technology and the U.S. economy." *American Economic Review*, 91 (1), 1–32. [235]

Keller, W. (1998), "Are international spillovers trade related? Analyzing spillovers among randomly attached trade partners." *European Economic Review*, 42, 1469–1481. [215]

Keller, W. (2002), "Geographic localization of international technology diffusion." *American Economic Review*, 92 (1), 120–142. [212, 215, 216, 235]

Khan, S. and E. Tamer (2010), "Irregular identification, support conditions, and inverse weight estimation." *Econometrica*, 78, 2021–2042. [240]

Kim, E. H., A. Morse, and L. Zingales (2006), "Are elite universities losing their competitive edge?" Working Paper 12245, NBER. [216, 236]

Kortum, S. and J. Lerner (1999), "What is behind the recent surge in patenting?" *Research Policy*, 28 (1), 1–22. [222]

Leamer, E. (2007), "A flat world, a level playing field, a small world or none of the above?" *Journal of Economic Literature*, XLV (1), 83–126. [212]

Lee, S. (2008), "Estimating panel data duration models with censored data." *Econometric Theory*, 24, 1254–1276. [218, 240]

Marshall, A. (1890), *Principles of Economics*. Macmillan, London. [211]

Moretti, E. (2010), "Real wage inequality." Discussion Paper 3706, IZA. Available at http://emlab.berkeley.edu/~moretti/inequality.pdf. [235]

Pakes, A. (1986), "Patents as options: Some estimates of the value of holding European Patent stocks." *Econometrica*, 54 (4), 755–784. [216]

Peri, G. (2005), "Determinants of knowledge flows and the effect on innovation." *Review of Economics and Statistics*, 87 (2), 308–322. [228]

Redding, S. (2009), "Economic geography: A review of the theoretical and empirical literature." Discussion Paper 904, Centre for Economic Performance. [216]

Ridder, G. and I. Tunali (1999), "Stratified partial likelihood estimation." *Journal of Econometrics*, 92, 193–232. [216, 218, 219]

Thompson, P. (2006), "Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations." *Review of Economics and Statistics*, 88 (2), 383–388. [212, 215, 216, 219, 242]

Thompson, P. and M. Fox-Kean (2005), "Patent citations and the geography of knowledge spillovers: A re-assessment." *American Economic Review*, 95 (1), 450–460. [215, 220, 227]

Wooldridge, J. M. (2007), "Inverse probability weighted estimation for general missing data problems." *Journal of Econometrics*, 141 (2), 1281–1301. [220]

Zucker, L., M. Darby, and M. Brewer (1998), "Intellectual property and the birth of US biotechnology enterprises." *American Economic Review*, 88 (1), 290–306. [232]

Submitted March, 2010. Final version accepted May, 2011.