CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 818

August 2007

Is Distance Dying at Last? Falling Home Bias in Fixed Effects Models of Patent Citations

Rachel Griffith, Sokbae Lee and John Van Reenen

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
We examine the "home bias" of international knowledge spillovers as measured by the speed of patent citations (i.e. knowledge spreads slowly over international boundaries). We present the first compelling 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 citations between 1975 and 1999. Home bias declines substantially when we control for fixed effects: there is practically no home bias for the more "modern" sectors such as pharmaceuticals and information/communication technologies.

Keywords: Fixed effects, home bias, patent citations, knowledge spillovers
JEL Classifications: O32, O33, F23

This paper is published as part of the Centre's Productivity and Innovation Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
This paper is produced as part of the project Science, Innovation, Firms and Markets in a Globalised World (SCIFI-GLOW), a Collaborative Project funded by the European Commission's Seventh Research Framework Programme, Contract Number SSH7-CT-2008-217436. Any opinions expressed here are those of the author(s) and not those of the European Commission.

We would also like to thank the Economic and Social Research Council for its financial support for the Institute for Fiscal Studies and the Centre for Economic Performance and the Leverhulme Trust for its financial support for the Centre for Microdata Methods and Practice. Sharon Belenzon, Richard Freeman, Jerry Hausman, Bruce Meyer and seminar participants at Berkeley and UCL/IFS have offered several useful comments.

Rachel Griffith is a Deputy Research Director at the Institute for Fiscal Studies, London. She is also a Professor of Industrial Organisation at University College London, a Research Associate of CEPR and a Fellow at AIM. Email: r.griffith@ifs.org.uk. Sokbae Lee is a lecturer with the Department of Economics, University College London. John Van Reenen is the Director of the Centre for Economic Performance, London School of Economics and a Research Fellow at CEPR.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 978 0 85328 075 0

"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)

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 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" (Friedman, 2005, Cairncross, 1997; Coyle, 1997) 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 geography plays little role.

There are several counter-arguments 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, 2006 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 will 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 we look at all successful applications to the US 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 14% slower in citing Germans patents than Germans themselves and the French were about 4% 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, British quicker at citing other British, etc. 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 5% slower in citing Germans and the French are only about 1% slower in citing Germans, than the Germans themselves. Table 1 holds the underlying data: the average time to the first citation in the early period from a German inventor to another German inventor was 1,559 days compared to 1,770 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 966 days. But the fall was even greater elsewhere: the time to first American citation fell to 1,016 (a fall of 754 days compared to a fall of 593 days for a German to German first citation).

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.

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, 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 domestic economy.[3] This approach has the advantage of using a direct measure of technology. But it shares two general problems of the R&D spillover literature. First, it is necessary to identify the

[1] There are other interesting features in Figure 1 over and above the general fall in home bias. First, Japanese inventors appear particularly quick at citing other countries' inventors and this has grown stronger over time. Second, although home bias has fallen for the US with respect to the Rest of World, it appears to have increased with respect to the main EU countries (Germany, France and Britain). As we will see in the econometric section, once we control for other factors, there is not much evidence for home bias of US inventions in the later period.

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

[3] Note that the production function is augmented to control for the firm's own R&D. 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). The best work is at the firm level which finds evidence that countries' behind the frontier benefit much more from frontier R&D than vice versa. See Bransetter (2001), Bransetter and Sakibora (2002) and Griffith, Harrison and Van Reenen (2006).

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 second, 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 which helped the second innovation. A classic paper in this field is Jaffe, Trajtenberg and Henderson (1993), which uses 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) when 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].

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 (e.g. the US and Japan) and assume that higher quality patents will be cited more quickly than lower quality patents. If US inventors produce, on average, higher quality patents and US inventors are more nimble at using the ideas of other countries we will observe that US based inventors tend to cite other US inventors more than they cite Japanese inventors. This will give the impression of "home bias" whereas in fact it is to do with the higher average quality of US inventors in the generation of new knowledge (and absorption of older knowledge). Controlling for fixed effects will therefore reduce the degree of "home

[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) argue that using more disaggregated patent classes drives away localization effects within the US, but they still observe home bias between the US and other countries.See also Henderson, Jaffe and Trajtenberg (2005) for a rejoinder.

bias" observed in naive estimators.

Now consider a second scenario where inventions in Japan remain of lower quality on average than in the US, but Japanese inventors are faster at absorbing old knowledge than their US counterparts. This will make it appear that Japanese inventors cite US 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 non-fixed effects estimators. In conclusion, the fixed effects bias could go in either direction, but certainly could be important.

Using a duration model without fixed effects we find evidence of large home bias, in line with the current 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). In the more "modern" technological sectors (such as the ICT, Information and Communication Technologies sector and the Pharmaceutical sector), where communication costs are likely to be relatively low, we find no systematic evidence for home bias. In the more "traditional" sectors (such as Chemicals and Mechanical Engineering), however, we do find evidence for systematic home bias, even after controlling for fixed effects. Our most important finding is that even after controlling for fixed effects, other covariates and censoring inter-country home bias appears to have *fallen* over time. This is consistent with the raw data shown in Figure 1 and Table 1. The only other econometric evidence that we are aware of that shows that geography matters less over time is the lower apparent degree of spillovers within elite US university departments (Kim, Morse and Zingales, 2006)[5]. By contrast, our work covers the entire economy.

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

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 in Section 5.

2 Econometric Modelling Strategy

Our approach to estimate the impact of home bias on knowledge spillovers is based on 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 distances, and researchers with direct knowledge of the invention become more geographically disperse.

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.[6]

Let subscript i index cited patents and subscript j citing patents. Under this convention, let $\tilde{Y}_{ij}$ denote the j-th citation duration for the i-th patent, that is the number of days from the date when the i-th cited patent is granted to the date when the j-th citing patent is granted, where $i = 1, \ldots, n$ and $j = 1, \ldots, J$. Here n is the 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 j-th citing patent for the i-th cited patent and U_i

[6]See, for example, Chamberlain (1985), Ridder and Tunalı (1999), Horowitz and Lee (2004), and Lee (2007).

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 $\tilde{Y}_{ij} = \tilde{y}_{ij}$ conditional on $X_{ij} = x_{ij}$ and $U_i = u_i$ has the form

$$\lambda_i(\tilde{y}_{ij}) \exp(x'_{ij}\beta + u_i), \tag{1}$$

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

The citation durations $\tilde{Y}_{ij}$ 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 j-th citing patent for the i-th cited patent. We allow U_i to be arbitrarily correlated with X_{ij} and do not impose any distributional assumptions on U_i, and 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.

Under the conditional independence assumption, such that $\tilde{Y}_{ij}$ are independent of each other conditional on (X_{ij}, U_i), we can estimate β using a conditional likelihood approach (e.g. Ridder and Tunali,1999). The idea behind the conditional likelihood approach is as follows. Assume that there are only two potential citing patents $(J = 2)$ and let (1) denote a random variable that indicates which of the two citation durations is first. The probability that the observed first citation duration is first, conditional on the duration of the first citation, is given by

$$\begin{aligned}
&\Pr[(1) = 1|\tilde{Y}_{i(1)} = \tilde{y}_{i1}, X_{i1} = x_{i1}, X_{i2} = x_{i2}, U_i = u_i] \\
&= \frac{\lambda_i(\tilde{y}_{i1}) \exp(x_{i1}'\beta + u_i)}{\lambda_i(\tilde{y}_{i1}) \exp(x_{i1}'\beta + u_i) + \lambda_i(\tilde{y}_{i1}) \exp(x_{i2}'\beta + u_i)} \\
&= \frac{\exp(x_{i1}'\beta)}{\exp(x_{i1}'\beta) + \exp(x_{i2}'\beta)}, \qquad (2)
\end{aligned}$$

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.

A usual problem with analyzing such data is censoring. Given any dataset there will be some patents that have not (yet) been cited, but which could in 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 31st 1999). We therefore treat the censoring as independent of citation durations and covariates, 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 the Appendix for details of our estimation method.

There are two main differences between our approach and the more usual Jaffe and Trajtenberg (1999) approach. First, we control for unobserved heterogeneity in a way that they cannot. Second, we use only a sub-sample 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.[7]

3 Data

To implement this estimator we use data from the NBER US Patent Citations Data File.[8] 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.[9] We focus on inventors located in the G5 countries - US, Japan, France, Germany and Great Britain. We group the remaining EU countries together,[10] and then consider the Rest of the World ("RW") as the residual category. Unsurprisingly, the US is the leading country with nearly 1.2 million patents, and Japan is second with nearly 400,000. We split our sample into two sub-periods, 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

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

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

[9]Where there is more than one inventor we follow Jaffe et al (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.

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

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[11]. For example, in our data (see Table 3), for 26% of patents in the chemicals technology sector we never see a citation, for 15% we see only one citation, and for the remaining 59% 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 or not the citation is a self-citation (i.e. whether the assignee is the same on the cited and citing patent) and whether or not 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 sub-category. Table 3 reports some summary statistics for these control variables by technology category. In Chemicals, 24% of all first citations are self-citations, and this falls to 20% for the second citation. On average across technology sectors about 20% are self citations and this declines by 2-4% from the first citation to the second citation. More than 60% of citations are from the same technology class, and the average number of patents in the citing country and technology sub-category varies from 24,300 to 49,600. 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 the seven cited regions and six technology

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

categories, and allow all the coefficients to vary across these groups.

4.1.1 An example - Mechanical 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 Mechanical Engineering in Great Britain. Each column in Table 4 reports the results from a different regression. The omitted category is the own country - the location of the cited patent - which in this case is Great Britain (GB). There are potentially 13,951 cited patents in Mechanical Engineering in Great Britain over this time period, and from these 8,482 patents are cited at least twice. 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.

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), i.e. compared to equation (1) we assume,

$$\lambda(\tilde{y}_{ij}) \exp(x'_{ij}\beta). \tag{3}$$

To keep the sample the same as when estimating 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 British inventors in mechanical engineering faster (a positive coefficient) or slower (a negative coefficient) than inventors from the omitted category (which is always the own country, in this case Great Britain). If there is home bias we expect *negative* coefficients on the other country dummies, i.e. they are slower to cite than home inventors. In column (1) we see negative and significant coefficients on all country dummies; these suggest strong support for home bias. Japanese inventors are the swiftest foreign group to cite British inventors - they are only 13% slower than British inventors themselves. By contrast inventors in the US are 31% slower to cite British 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 with the first two citation durations (without correcting for censoring).[12] When fixed effects are included most coefficients become closer to zero and the French and EU dummies become statistically insignificantly different from zero. On average, this suggests that failure to control for unobserved heterogeneity increases the degree of home bias. Note, however, that the coefficient on Japan becomes larger in absolute magnitude, suggesting that the fixed effects bias for Japan was to reduce the degree of home bias. This reinforces the point that the direction of bias from failure to control for fixed effects cannot be signed *a priori*. The simple fixed effects estimator in column (2) ignores the problem of censoring. In column (3) we also allow for censoring, which raises the absolute magnitude of most of the coefficients (and increases the standard errors), but 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 27% faster than non self-citations in column (3)). Similarly, patents in the same technology class cite each other significantly faster (12% faster than patents in different technology classes according to column (3)).

We continue to illustrate the method by looking across all countries but still restricting ourselves to patents in the Mechanical Engineering category. In Table 5 each row contains parameter estimates from a separate multiple spell duration model for each country. For example, the third row shows the results from column (3) of Table 4 (the

[12]Specifically, the estimator maximizes the likelihood equation (A1) in the Appendix without the correction term $G_n(\max[Y_{i1}, Y_{i2}])$.

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"), i.e. 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 all countries are slower to cite French inventors than the French themselves: the coefficients for German and British inventors are only weakly significant, other European inventors are 23% slower to cite French inventors, Japanese are 19% slower, the Americans 32% slower and the rest of the world 52% slower. So, just as in the British case, we see substantial evidence of apparent home bias. Note that all regressions include unreported controls for whether the citation is a self citation, whether it is in the same technology sub-category (three digit) and the total number of citing patents in that country and technology class ("base"). Most of these controls are highly significant and would lead to the impression of even more substantial home bias if omitted.

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 significance at the 5% level, a small circle significance at the 10% level, and a tiny circle *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 or displayed.[13] 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. Britain, for example, has a full

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

row of large circles indicating significant home bias, whereas the United States does not (consistent with 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 line box), compared to between the non-EU countries and EU countries - the top right quadrant contains many rejections suggesting that Japan, the US and other countries are much slower in citing European patents than their own inventors. Furthermore, it can be seen that there exists an interesting asymmetry between the European block and the Japan/US block, in the sense that European inventors are quick to cite Japanese and American patents but Japan/US-located inventors are slow to cite European patents.

Note that here, and in a few other cases, we obtain positive and significant coefficients on US-cited patents. These suggest that, for example, French (FR) and British (GB) inventors are quicker to cite US patents than are US inventors themselves. This may reflect the fact that the patents we are focusing on are patents taken out in the US patent office, suggesting that there will be a bias of all countries towards citing American inventors.

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[14] for no home bias, we reject at the 5% level in more than 8 out of 10 cases (81%). This is consistent with evidence from the analysis of citations data in other econometric studies (Jaffe and Trajtenberg, 1999; Jaffe et al, 2005; Thompson and Fox-Kean, 2005). However, the picture changes when we control for unobserved heterogeneity (columns (4) through (6)). Comparing column (5) to column (2), for example, the rejection rate (the proportion of possible rejections that are in fact rejected) falls from 81% to 36%. 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) through (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 dramatically when moving from the no-fixed-effects specifications to the preferred specifications.

A second feature of Table 6 and Figure 3 is that there is a distinct pattern across technology sectors in terms of home bias. In particular, there is less evidence of home bias for the "modern" sectors of "Drugs and Medical", "Electrical and Electronic", and "Computers and Communications" (ICT) once we control for unobserved heterogeneity - i.e. we see few rejections of home bias in columns (4)-(9) in Table 6.[15] Whereas we reject home bias for the "more traditional" sectors of Mechanical Engineering and Chemicals more often in Table 6. This is consistent with the idea that knowledge spreads

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

[15]For example, in column (5) there are eight rejections in computers and eleven rejections in drugs. By comparison there were twenty rejections in chemicals and eighteen in mechanical.

much more quickly in the high tech, modern sectors of Computers and Communications and Drugs & Medicines than in the older sectors. It is also congruent with Peri (2004) who finds that knowledge spreads much more quickly across regional boundaries in the computer and communication sector.

A third feature of Figure 3 is that the Rest of World (mainly non-OECD countries) are much slower in citing the patents of the OECD countries. This suggests that non-OECD countries are more "cut-off" from international pools of knowledge, either by dint of their distance, infrastructure or 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" period (1990-1999) where there are a similar absolute number of citations in each period (see Table 2). We re-estimate all of our models on these two sub-periods separately. We report a summary of these results in Tables 7 and 8 and Figure 4.[16] It is particularly 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.

We start in Table 7 by reporting results using the no fixed effect hazard model estimator. In columns (1) and (2) we see that there is a large decline in rejection rates across all technology sectors over time. No home bias is rejected in 70% of cases in the early period, but only for 38% 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 (i.e. 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). There is substantial home bias

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

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 in the traditional sectors remains dramatic. On average the rejection rate falls from 66% to 33%, but the fall is particularly strong in Chemicals (from 21 rejections in the early period to only 6 in the late period). 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 large decline in home bias.

As we saw above, controlling for unobserved heterogeneity is important. In Table 8 we find that in both periods, the level of home bias is lower when we control for fixed effects (and censoring), but the trend reduction in home bias over time is just as apparent in the fixed effects models as it was in the no fixed effects models of Table 7. Looking at the first two columns of Table 8 there is a fairly clear pattern. For the three "traditional" sectors (Chemicals, Mechanical Engineering and Others) it appears that home bias has fallen substantially. For Chemicals the number of rejections fell from 16 in the early period (pre 1989) to 8 in the 1990s (i.e. the rejection rate fell from 38% to 19%). For Mechanical Engineering the fall was from 20 (48%) to 6 (14%) and for the other industries it fell from 20 (48%) to 11 (26%). These are substantial declines, suggesting a big fall in home bias over time for the traditional sectors. By contrast, the "modern" sectors have seen little change, mainly because they had much less home bias to begin with. We actually see an increase in the number of rejections from 2 to 5 in Computers & Communications and a constant number of 8 in Drugs & Medicines. These are the two sectors where there has been the most discussion of "clustering" (e.g. ICT in Silicon Valley and biotechnology in Cambridge, Massachusetts)[17] so it may be unsurprising that the forces of agglomeration remain unchanged in these sectors. On

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

the other hand, it is worth noting that there was very little evidence of home bias once we have controlled properly for unobserved heterogeneity, so that the absolute number of rejections is still very low, even in the late period. These results are also shown in Figure 4, where the left hand side diagrams are of the early period and the right hand side diagrams are of the late period: the later period has far fewer "circles" (evidence for home bias) than the earlier period.

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 appears to be the first concrete quantitative evidence for an aspect of globalization that is much discussed but never proven - the increasing propensity of knowledge to slip over geographic boundaries. It 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 Robustness of results

Could there be reasons why the apparent decline in home bias is spurious? Firstly, 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, so this cannot be the reason. Secondly, 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 the US, Japan and the Rest of the World are citing European patents more quickly (relative to their own patents). This can be seen from the decline in large circles in the upper right hand quadrants of the "traditional industries". But this ignores the fact that there has been some decline in home bias *within* European countries. More importantly, 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). Thirdly, we also tried using different cut-off years and found that this lead to similar results. For example, we obtain qualitatively similar results using 1985 as a cut-off year. For example, in chemicals the number of rejections using all countries decreases from 21 in the pre-1985 period to 10 in the post-1985 period. For mechanical engineering the fall was from 16 to 9.[18]

We have used only the first two citations to measure home bias. Why not use the third, fourth, fifth, etc. 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. So the first few citations are exactly what we are interested in. In addition, we have argued for a smaller number of citations on 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 is 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 $J = 3$, i.e. 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 estimator exactly the same way as in equation (2) by replacing the subscripts 1 and 2 with subscript 2 and 3, respectively.[19]

[18] As before, the "modern" sectors have seen an increase in the number of rejections from 7 to 11 in pharmaceuticals and from 5 to 9 in electrical and electronic.

[19] Similarly, if $J = 4$, one can show that the probability that the observed third citation is third,

Our qualitative findings did not change. For example, in Table 6, for the 5% level, the number of rejections falls from 203 to 94 as we control for unobserved heterogeneity of citing patents. When we use the second and third citations, for the same level, the number of rejections changes from 202 to 87; when we use the third and fourth citations, the number falls from 192 to 42. 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.

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 database of over two million citations between 1975 and 1999. We have focused 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

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.

of cases, whereas our preferred models indicate home bias in only a minority of cases. Second, home bias is much stronger in the "traditional sectors" (such as chemicals and mechanical engineering) than in more "modern" sectors (such as computing), consistent with the idea that information diffuses faster internationally in these sectors. Finally, 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.

References

Belenzon, Sharon and John 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

Bernstein, Jeff and Pierre Mohnen (1998), "International R&D spillovers between US and Japanese R&D intensive sectors" *Journal of International Economics*, 44, pp. 315-38

Bloom, Nick and John Van Reenen (2002), "Real options, patents, productivity and market value: Evidence from a panel of British firms", *Economic Journal*, 112, 478, C97-C116

Branstetter, Lee (2001), "Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the U.S. and Japan", *Journal of International Economics*, 53, pp. 53-79

Branstetter, Lee and Mariko Sakakibara (2002) "When do research consortia work well and why? Evidence from Japanese panel data" *American Economic Review,* 92,1, pp. 143-159

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

Chamberlain, Gary (1985) "Heterogeneity, omitted variable bias, and duration dependence" in: J.J. Heckman and B. Singer, (eds.) *Longitudinal analysis of labor market data* Cambridge: Cambridge University Press.

Coe, David and Elhanen Helpman (1995) "International R&D Spillovers" *European Economic Review*, 39 pp. 859-887

Coyle, Diane (1997) *Weightless World*, Princeton: Princeton University Press

Disdier, Anna-Celia and Keith Head (2005) "The Puzzling persistence of the Distance Effect on bilateral trade" forthcoming *Review of Economics and Statistics*

Evans, Carolyn and James Harrigan (2005) "Distance, time and specialization: Lean Retailing in general equilibrium" *American Economic Review*, 95, pp. 292-313

Friedman, Thomas (2005) *The World is Flat*, New York: Farrar, Strauss and Giroux

Griffith, Rachel, Steven Redding, and John Van Reenen (2004) "Mapping the The Two Faces of R&D" *Review of Economics and Statistics*, 86 (4) pp. 883 - 895.

Griffith, Rachel, Rupert Harrison, and John 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*, 76 (5), pp. 1859-1875

Griliches, Zvi (1992), "The search for R&D spillovers", *Scandinavian Journal of Economics* 94, pp. S29-S47

Hall, Bronwyn, Adam Jaffe and Manuel Trajtenberg (2001) "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools" NBER Working Paper 8498; see also http://www.nber.org/patents/

Hall, Bronwyn, Jaffe, Adam and Manuel Trajtenberg (2005), "Market Value and Patent Citations: A First Look" *RAND Journal of Economics,* 36, pp.16-38

Horowitz, Joel L., and Sokbae Lee (2004) "Semiparametric estimation of a panel data proportional hazards model with fixed effects", *Journal of Econometrics*, 119, pp. 155-198.

Henderson, Rebecca, Adam Jaffe and Manuel Trajtenberg (2005), "Patent citations and the geography of spillovers: A Reassessment - Comment" *American Economic Review* 95, 1, pp. 461-464

Head, Keith, Mayer, Thierry and Ries, John (2007) "How remote is the offshoring threat?", mimeo, Univeristy of Toronto

Jaffe, Adam (1986), "Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits and market value", *American Economic Review* 76, pp 984-1001

Jaffe, Adam and Manuel Trajtenberg (2002) *Patents, citations and innovation: A Window on the knowledge economy*" Cambridge: MIT Press

Jaffe, Adam and Manuel Trajtenberg (1999), "International knowledge flows: Evidence from patent citations", *Economics of Innovation and New Technology*, 8(1-2), pp 105-36

Jaffe, Adam, Manuel Trajtenberg and Rebecca Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations", *Quarterly Journal of Economics* 108 (3), pp 577-598

Keller, Wolfgang (1998) "Are international spillovers trade related? Analyzing spillovers among randomly attached trade partners" *European Economic Review*, 42, pp. 1469-1481

Kim, E Han, Adair Morse and Luigi Zingales (2006) "Are Elite Universities Losing their Competitive Edge?" NBER Working Paper 12245

Leamer, Edward (2007) "A Flat world, a level playing field, a small world or none of the above?" *Journal of Economic Literature*, XLV(1), 83-126

Lee, Sokbae (2007) "Estimating Panel Data Duration Models with Censored Data" *Econometric Theory*, forthcoming.

Marshall, Alfred (1890) *Principles of Economics*, London: Macmillan

Pakes, Ariel (1986) "Patents as Options: Some estimates of the value of holding European Patent stocks" *Econometrica,* 54(4), pp755-784

Peri, Giovani (2004) "Determinants of Knowledge Flows and the effect on innovation" *Review of Economics and Statistics*, 87(2), pp. 308-322

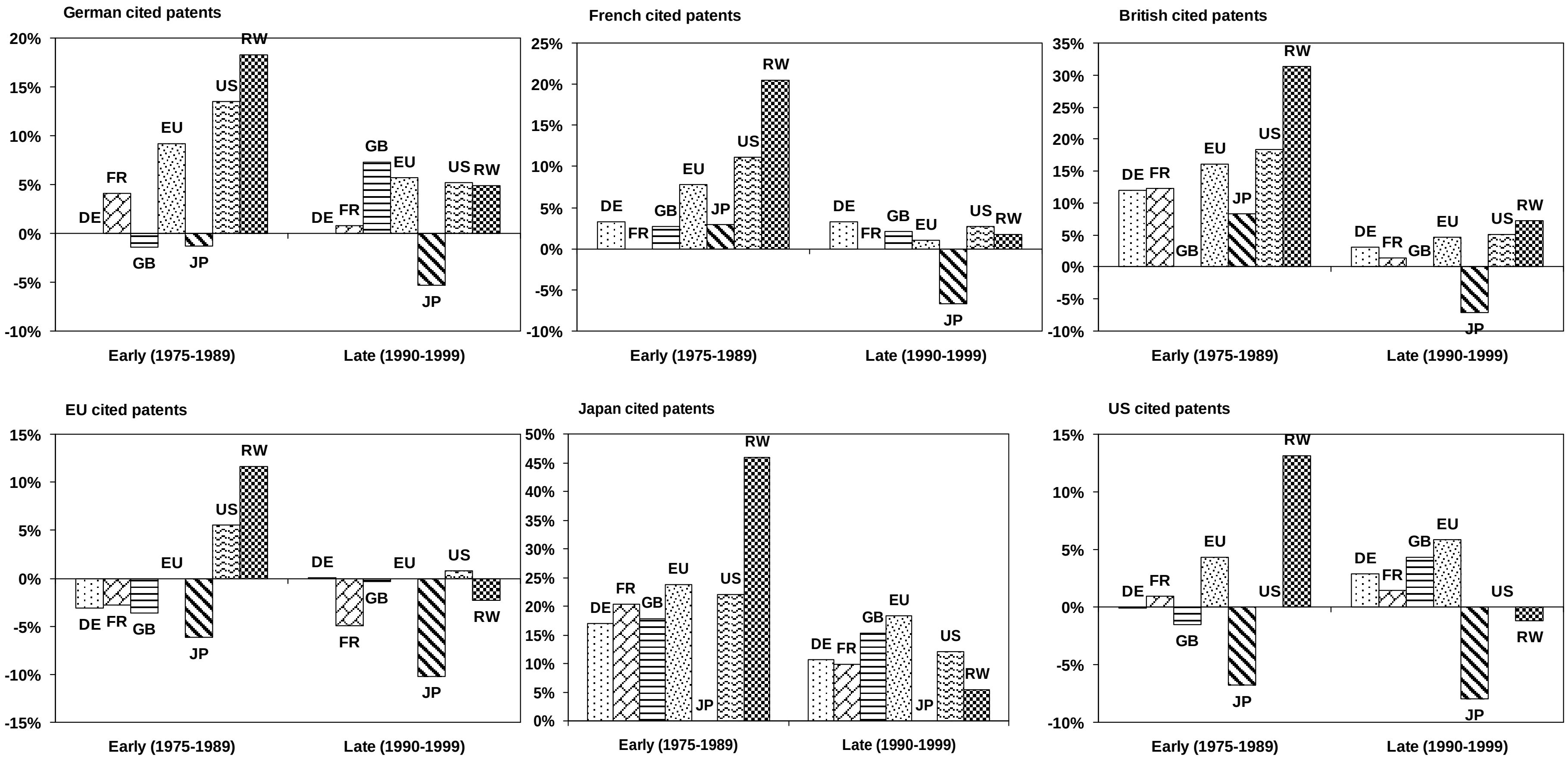
Ridder, Geert, and Insan Tunali (1999) "Stratified partial likelihood estimation" *Journal of Econometrics*, 92, pp. 193-232.

Thompson, Peter and Melanie Fox-Kean (2005) "Patent citations and the geography of knowledge spillovers: A re-assessment" *American Economic Review*, 95 (1) pp. 450-461

Wooldridge, Jeffrey M. (2007), "Inverse probability weighted estimation for general missing data problems" *Journal of Econometrics*, forthcoming

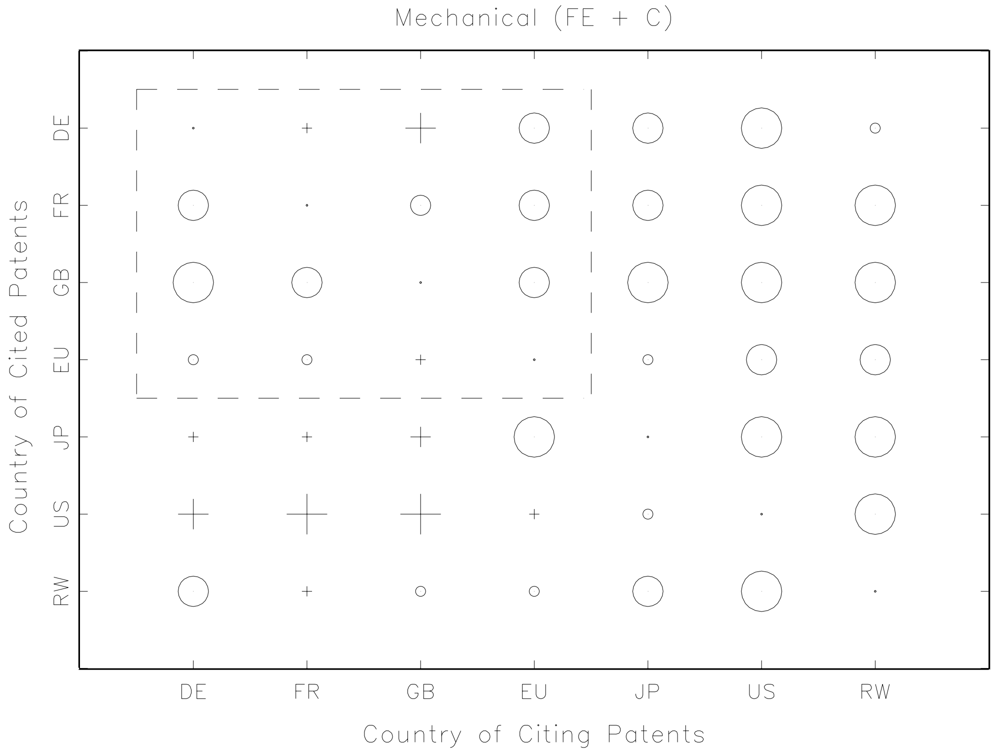
Zucker, Lynne, Michael Darby and Marilyn Brewer (1998) "Intellectual Property and the Birth of US Biotechnology Enterprises" *American Economic Review,* 88(1), pp. 290-306

Figure 1: Time to first citation, by cited and citing inventor location



Notes: This graph shows the relative time (in mean number of day) from the date that a Germany inventor was granted a patent until the first citation of that patent, by the location of the inventor that 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 4% longer than when the first citation was made by a Germany inventor (i.e. the mean citation length to a German inventor was 1559 days compared to 1623 days (1623=1599*1.04) to a French inventor).
Table 1 shows the raw numbers for all cells.

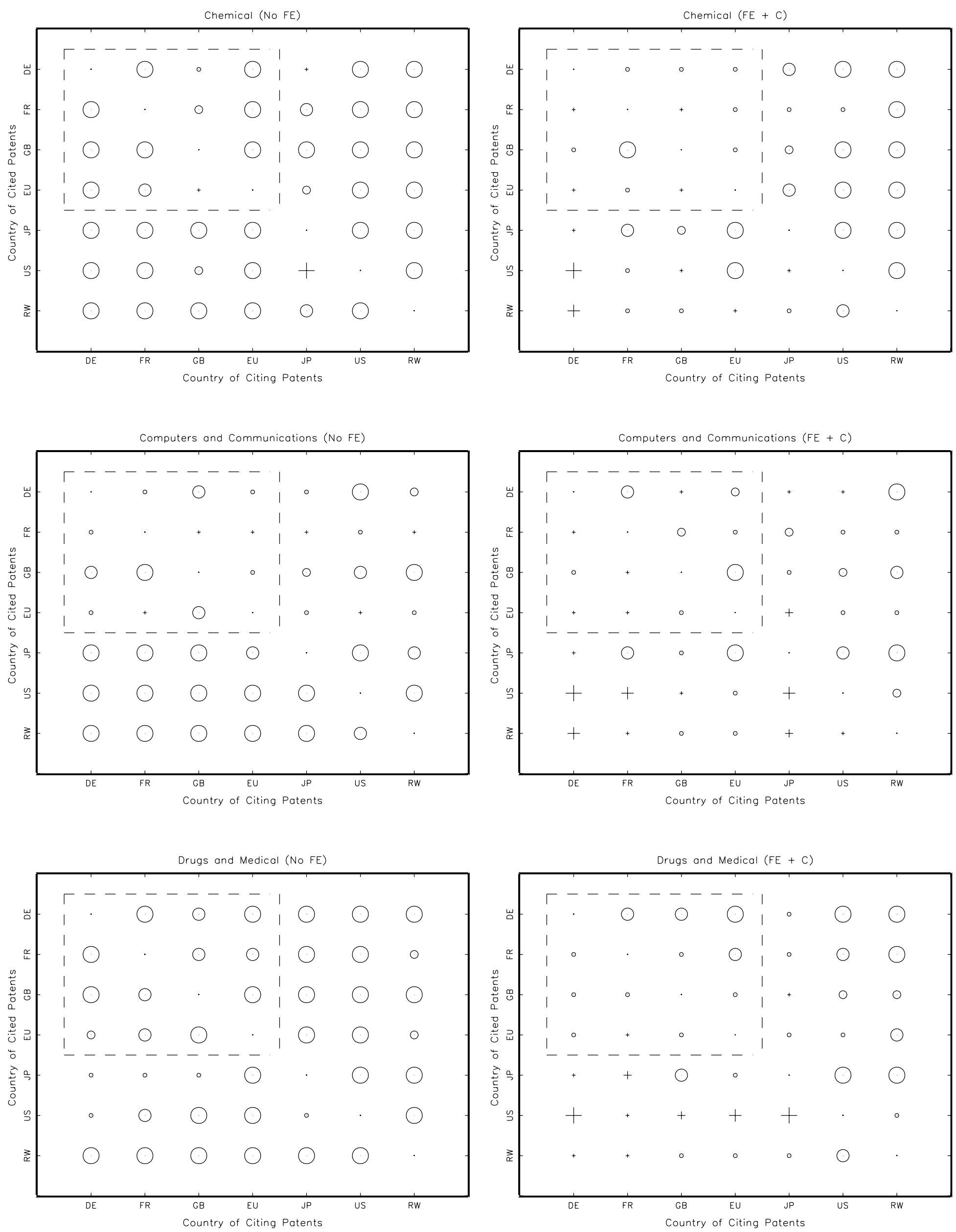
Figure 2: Graphical Representation of Estimation Results



Notes: Each cell in Figure 2 corresponds to the equivalent cell in 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-side test for the null hypothesis that the corresponding coefficient is zero. A large circle represents significance at the 1% level, a medium circle significance at the 5% level, a small circle significance at the 10% level, and a tiny circle *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.

Figure 3: No Fixed Effects ("No FE") and Fixed Effects with Censoring ("FE+C")

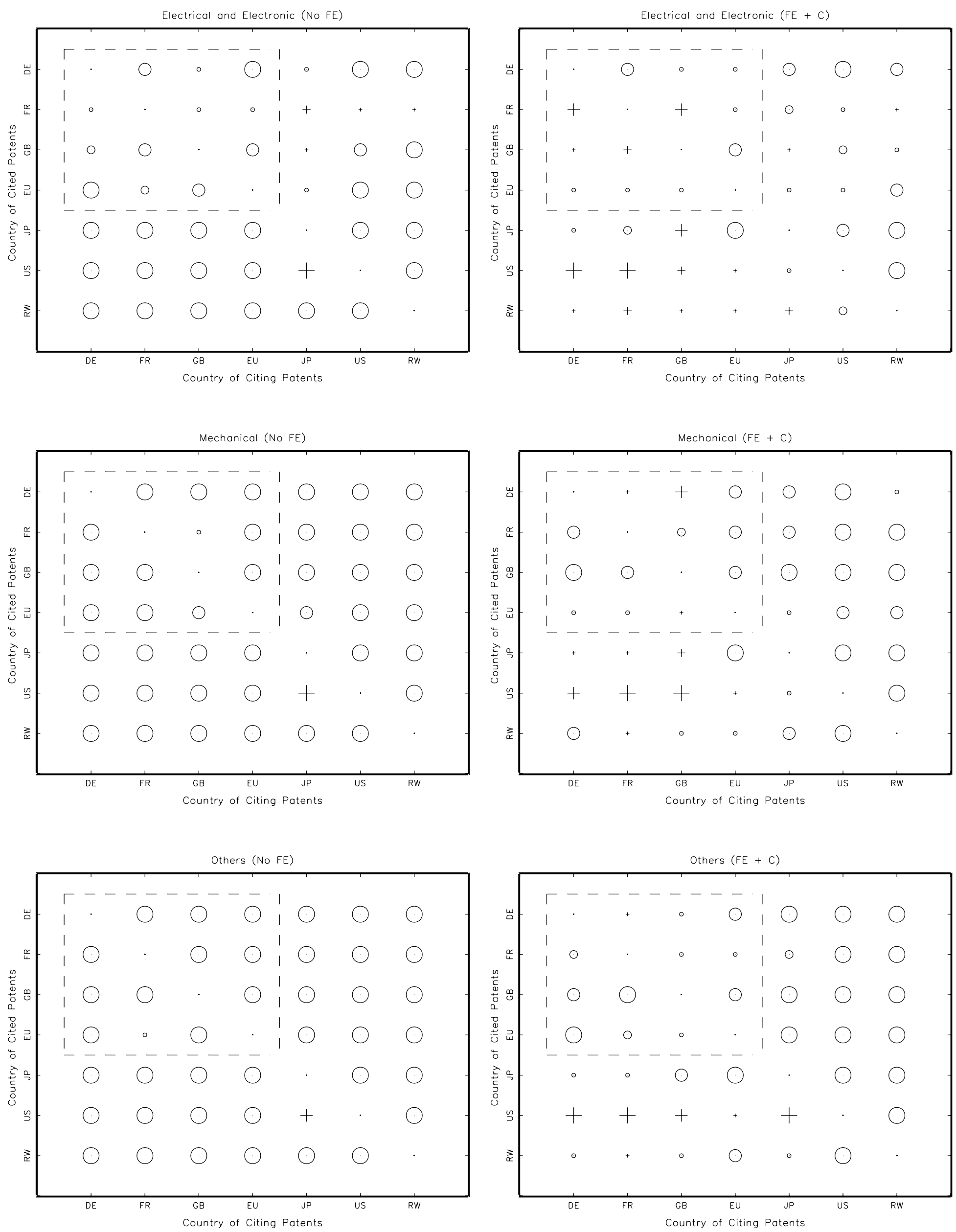
Panel A: No Fixed Effects *Panel B: Censored Fixed Effects*



Notes: 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 quadrant with dashed lines contains the cross-citations from the European Countries.

Figure 3: No Fixed Effects ("No FE") and Fixed Effects with Censoring ("FE+C") (Continued)

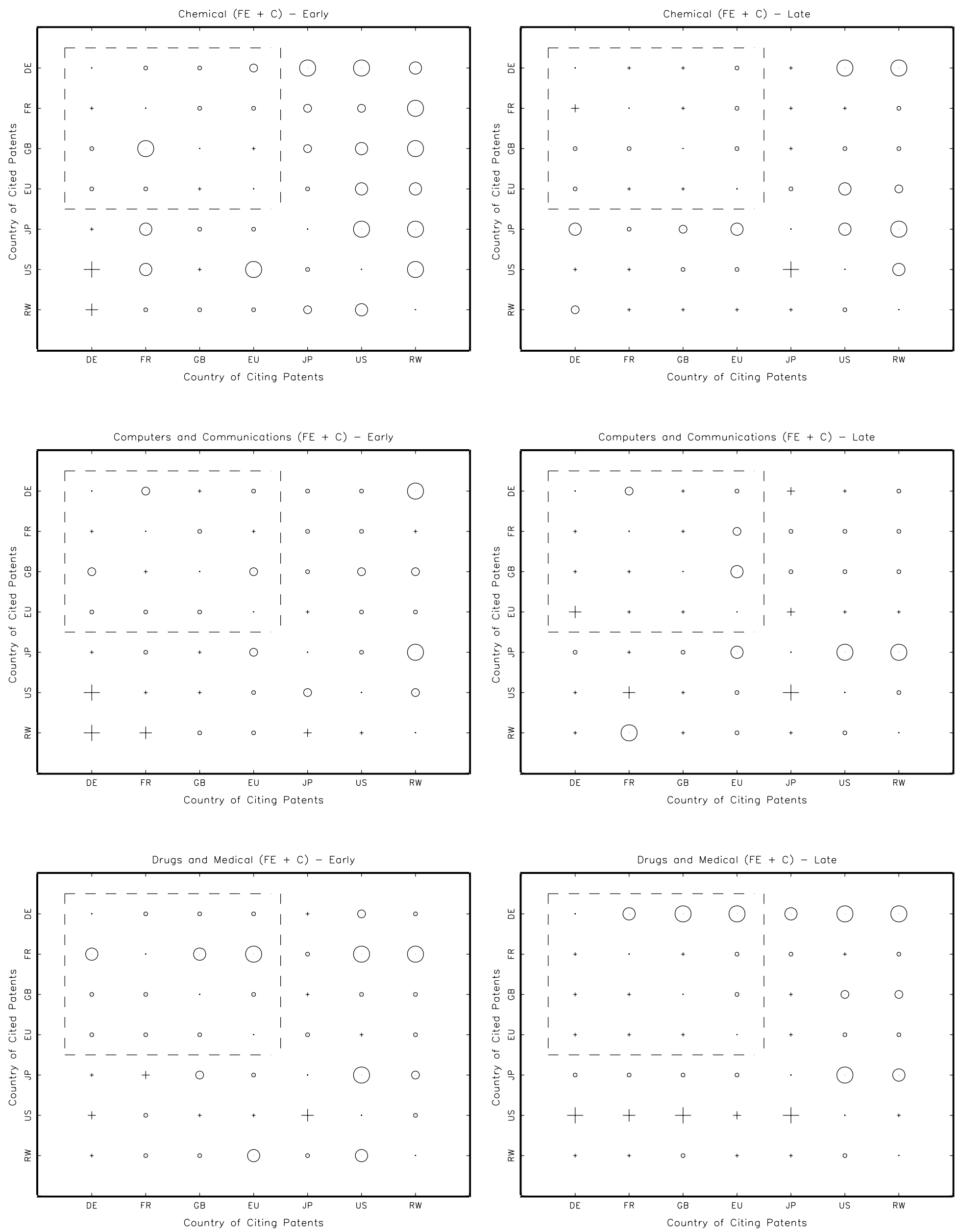
Panel A: No Fixed Effects *Panel B: Censored Fixed Effects*



Notes: 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 quadrant with dashed lines contains the cross-citations from the European Countries.

Figure 4: Early Period vs. Late Period

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



Notes: The left-hand side diagrams are estimation results for the early period (1975-1989) and the right hand side diagrams are for the late period (1990-1999). Estimation results are from our preferred fixed effects plus censoring specifications. The upper left quadrant with dashed lines contains the cross-citations from the European Countries.

Figure 4: Early Period vs. Late Period (Continued)

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

Notes: The left-hand side diagrams are estimation results for the early period (1975-1989) and the right hand side diagrams are for the late period (1990-1999). Estimation results are from our preferred fixed effects plus censoring specifications. The upper left quadrant with dashed lines contains the cross-citations from the European Countries.

Table 1: Time to first citation, by cited and citing inventor location

Period: 1975 - 1989							
	Citing: DE	FR	GB	EU	JP	US	RW
Cited:							
DE	1559	1623	1537	1702	1539	1770	1844
FR	1606	1555	1596	1676	1600	1727	1874
GB	1644	1649	1469	1705	1590	1738	1930
EU	1638	1643	1629	1690	1586	1784	1886
JP	1382	1423	1392	1464	1183	1443	1726
US	1637	1654	1615	1710	1528	1639	1855
RW	1728	1753	1755	1799	1711	1814	1795

Period: 1990 - 1999							
	Citing: DE	FR	GB	EU	JP	US	RW
Cited:							
DE	966	973	1037	1021	915	1016	1013
FR	1015	983	1004	994	918	1009	1000
GB	962	946	933	976	866	980	1000
EU	995	945	991	995	893	1002	972
JP	879	872	915	940	794	889	838
US	900	887	912	925	804	874	864
RW	965	926	945	937	814	931	834

Notes: 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 that made the first citation. For example, the number in the top panel for the first French (FR) citation to a German (DE) patents 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 1623 days. The top and bottom panels show the average time to first citation for the period of 1975 to 1989 and that of 1990 to 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.

Table 2: Sample Sizes of Patent Citation Data

Technological Category	Period	Country of Cited Patents DE	FR	GB	EU	JP	US	RW	Total
Chemical	All	47076	13834	14635	22083	73463	229342	28680	429113
	Early	26876	7331	8924	11351	32499	129433	14878	231292
	Late	20200	6503	5711	10732	40964	99909	13802	197821
Computers and Communications	All	8556	6766	6347	7927	70874	133411	13154	247035
	Early	4110	3158	2764	2940	19862	45020	2871	80725
	Late	4446	3608	3583	4987	51012	88391	10283	166310
Drugs and Medical	All	12707	7061	8027	10101	18261	114125	13047	183329
	Early	5841	2764	3562	3427	6839	38390	4654	65477
	Late	6866	4297	4465	6674	11422	75735	8393	117852
Electrical and Electronic	All	25841	12045	10719	14114	85771	192136	26118	366744
	Early	14293	6388	6524	7378	30796	96573	8267	170219
	Late	11548	5657	4195	6736	54975	95563	17851	196525
Mechanical	All	46432	14059	13951	24393	96980	239476	32132	467423
	Early	26542	8253	9033	14078	42732	133172	15073	248883
	Late	19890	5806	4918	10315	54248	106304	17059	218540
Others	All	30229	11489	12300	21836	46424	283137	39542	444957
	Early	17565	6528	7519	12277	21313	151271	17664	234137
	Late	12664	4961	4781	9559	25111	131866	21878	210820
Total	All	170841	65254	65979	100454	391773	1191627	152673	2138601
	Early	95227	34422	38326	51451	154041	593859	63407	1030733
	Late	75614	30832	27653	49003	237732	597768	89266	1107868

Notes: 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: '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. 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.

Table 3: Summary Statistics for Patent Citation Data

Variable	Chemical	Computers and Communications	Drugs and Medical	Electrical and Electronic	Mechanical	Others
Proportion of patents with two or more citations	0.59	0.64	0.54	0.64	0.61	0.62
Proportion of patents with only one citation	0.15	0.11	0.13	0.13	0.15	0.15
Proportion of patents with no citation	0.26	0.25	0.33	0.23	0.24	0.23
Proportion of self-citation (first citation)	0.24	0.16	0.26	0.18	0.24	0.31
Proportion of self-citation (second citation)	0.20	0.14	0.22	0.15	0.21	0.28
Proportion of same technology class (first citation)	0.65	0.71	0.76	0.67	0.69	0.70
Proportion of same technology class (second citation)	0.63	0.71	0.76	0.66	0.68	0.69
Average of Base (first citation)	4.79	3.12	2.94	2.43	3.22	3.91
Average of Base (second citation)	4.96	3.15	3.04	2.49	3.30	3.97

Notes: 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 sub-category (1 unit = 10,000 patents).

Table 4: Estimation Results

Technological Category: Mechanical

Country of Cited Patents: Great Britain (GB)

Sample Size: 13951; Obs. with at least two citations: 8482

Variable	(1) No Fixed Effect	(2) Fixed Effect	(3) Fixed Effect plus Censoring
DE	-0.23	-0.17	-0.26
	(0.05)	(0.08)	(0.09)
FR	-0.21	-0.07	-0.20
	(0.07)	(0.11)	(0.12)
EU	-0.26	-0.07	-0.19
	(0.06)	(0.09)	(0.11)
JP	-0.13	-0.23	-0.30
	(0.05)	(0.08)	(0.09)
US	-0.31	-0.25	-0.35
	(0.05)	(0.08)	(0.09)
RW	-0.33	-0.24	-0.31
	(0.06)	(0.09)	(0.10)
Self Cit.	0.30	0.31	0.27
	(0.04)	(0.07)	(0.08)
Tech.Cat.	0.02	0.10	0.12
	(0.02)	(0.04)	(0.05)
Base	-0.07	0.06	0.04
	(0.06)	(0.09)	(0.12)

Notes: Standard errors are in the parentheses. The dummy variables for the location of an applicant of citing patent are 'DE' = Germany, 'FR' = France, '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 omitted category in citing patent country dummies is Great Britain (GB). The Self Citation and Technology Category variables are dummy variables. The Base variable is the number of patents in citing country and subcategory (one unit = 100,000 patents). Different columns show different estimates. Column (1) shows no-fixed-effect estimates using the only first citation duration, Column (2) shows fixed-effect (FE) estimates using first two citation durations, and Column (3) shows FE estimates accounting for censoring.

Table 5: Estimation Results of Mechanical (FE + C)

Mechanical (FE + C) - Full Sample

Country of Cited Patents	(1) DE	(2) FR	(3) GB	(4) EU	(5) JP	(6) US	(7) RW
			Country of Citing Patents				
DE		0.02	0.15	-0.12	-0.09	-0.22	-0.07
		(0.06)	(0.07)	(0.05)	(0.04)	(0.04)	(0.06)
FR	-0.18		-0.21	-0.23	-0.19	-0.32	-0.52
	(0.11)		(0.14)	(0.12)	(0.11)	(0.11)	(0.12)
GB	-0.26	-0.20		-0.19	-0.30	-0.35	-0.31
	(0.09)	(0.12)		(0.11)	(0.09)	(0.09)	(0.10)
EU	-0.05	0.00	0.02		-0.05	-0.12	-0.12
	(0.07)	(0.10)	(0.13)		(0.07)	(0.07)	(0.07)
JP	0.03	0.07	0.09	-0.16		-0.11	-0.31
	(0.03)	(0.07)	(0.06)	(0.05)		(0.03)	(0.05)
US	0.04	0.12	0.14	0.00	-0.02		-0.11
	(0.02)	(0.04)	(0.03)	(0.03)	(0.02)		(0.02)
RW	-0.15	0.03	-0.10	-0.08	-0.15	-0.22	
	(0.07)	(0.10)	(0.11)	(0.08)	(0.06)	(0.06)	

Notes: 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 a technology category called "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) through (7) correspond to the inventor location of the patent which subsequently cites the original patent. The left-out base country dummy is the cited patent's country. Country codes with corresponding country names are as follows: '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. In addition to country dummies, each hazard regression includes, as explanatory variables, dummy variables for self citation and technology sub-category and the number of patents in citing country and subcategory.

Table 6: Number of Rejections of No Home Bias using Entire Sample

Technological Category	Max. # of rejections	(1) No FE 10%	(2) No FE 5%	(3) No FE 1%	(4) FE 10%	(5) FE 5%	(6) FE 1%	(7) FE+C 10%	(8) FE+C 5%	(9) FE+C 1%
Chemical	42	38	35	32	23	20	14	19	17	13
Computers & Communications	42	27	25	18	9	8	3	13	8	4
Drugs & Medical	42	37	34	28	17	11	4	15	13	6
Electrical & Electronic	42	31	29	24	12	10	5	14	10	4
Mechanical	42	40	40	38	20	18	14	24	23	12
Others	42	40	40	40	28	23	22	26	23	18
Total	252	213	203	180	109	90	62	111	94	57
Percentage		0.85	0.81	0.71	0.43	0.36	0.25	0.44	0.37	0.23

Notes: 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 first citation duration, fixed-effect (FE) estimator using first two spells, and censored fixed effect (FE+C) estimator.

Table 7: Number of Rejections of No Home Bias using Sub-Samples

Estimation Method: No Fixed Effect Estimator

Technological Category	(1) All countries Early	(2) All countries Late	(3) OECD countries Early	(4) OECD countries Late	(5) EU countries Early	(6) EU countries Late
Chemical	33	9	21	6	8	0
Computers and Communications	23	10	14	7	3	1
Drugs and Medical	28	12	19	8	9	4
Electrical and Electronic	25	16	16	8	4	2
Mechanical	31	23	22	14	8	4
Others	37	26	26	16	11	5
Total	177	96	118	59	43	16
Max. # of rejections	252	252	180	180	72	72
Percentage	0.70	0.38	0.66	0.33	0.60	0.22

Notes: 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. The columns under “All countries” show the number of rejections for all coefficients for country dummies (42 coefficients), those under “OECD countries” 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 those under “EU countries” 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 8: Number of Rejections of No Home Bias using Sub-Samples

Estimation Method: Censored Fixed Effect Estimator

Technological Category	(1) All countries Early	(2) All countries Late	(3) OECD countries Early	(4) OECD countries Late	(5) EU countries Early	(6) EU countries Late
Chemical	16	8	9	5	1	0
Computers & Communications	2	5	0	3	0	1
Drugs & Medical	8	8	5	6	3	3
Electrical & Electronic	8	7	4	5	0	1
Mechanical	20	6	14	1	4	0
Others	20	11	13	5	3	0
Total	74	45	45	25	11	5
Max. # of rejections	252	252	180	180	72	72
Percentage	0.29	0.18	0.25	0.14	0.15	0.07

Notes: 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. The columns under "All countries" show the number of rejections for all coefficients for country dummies (42 coefficients), those under "OECD countries" 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 those under "EU countries" 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.

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 $(\tilde{Y}_{ij}, X_{ij}, U_i)$ 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 $\{(Y_{ij}, \Delta_{ij}, X_{ij}, C_i) : i = 1, \ldots, n, j = 1, \ldots, J\}$, where $Y_{ij} = \min(\tilde{Y}_{ij}, C_i)$ and $\Delta_{ij} = 1(\tilde{Y}_{ij} < C_i)$. 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 $J = 2$:

$$\begin{aligned} L(b) &= n^{-1} \sum_{i=1}^{n} \frac{\Delta_{i1}\Delta_{i2}}{G_n(\max[Y_{i1}, Y_{i2}])} \left\{ \left[1(Y_{i1} \leq Y_{i2}) \ln \left(\frac{\exp(X_{i1}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)} \right) \right] \right. \\ & \left. \quad + 1(Y_{i1} \geq Y_{i2}) \ln \left[\frac{\exp(X_{i2}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)} \right] \right\}, \end{aligned} \tag{A1}$$

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 (A1). Observe that the selection bias is corrected for by multiplying weights $G_n(\max\{Y_{i1}, Y_{i2}\})^{-1}$ in equation (A1). The reason why $G_n(\max\{Y_{i1}, Y_{i2}\})^{-1}$'s are proper weights is that

$$E\left[\frac{\Delta_{i1}\Delta_{i2}}{G(\max\{Y_{i1}, Y_{i2}\})} \middle| \tilde{Y}_{i1}, \tilde{Y}_{i2}, X_{i1}, X_{i2} \right] = 1 \tag{A2}$$

In other words, (A1) converges in probability uniformly over b to a limiting function to which an infeasible log-likelihood function would converge under no censoring. In maximizing (A1), 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 *(1) β is an interior point of a compact subset of $\mathbf{R}^d$ for some finite d. (2) The data $\{(Y_{i1}, Y_{i2}, X_{i1}, X_{i2}, \Delta_{i1}, \Delta_{i2}, C_i) : i = 1, \ldots, n\}$ are independent and identically distributed. (3) $\tilde{Y}_{i1}$ and $\tilde{Y}_{i2}$ are independent of each other conditional on (X_{i1}, X_{i2}, U_i). (4) $\lambda_i(\cdot)$ is strictly positive. (5) $E\left[\|X_{i1} - X_{i2}\|^2\right] < \infty$ and $E[(X_{i1} - X_{i2})(X_{i1} - X_{i2})']$ is nonsingular. (6) The censoring* *variable C_i is random with an unknown continuous probability distribution. (7) C_i is independent of $(\tilde{Y}_{i1}, \tilde{Y}_{i2}, X_{i1}, X_{i2}, U_i)$. (8) The survivor function of C_i, $G(c) \equiv Pr(C_i > c)$ is positive for every $c \in \mathbf{R}$.*

These assumptions are not unrestrictive, but in our application, they might be viewed as plausible. It is reasonable that the censoring time C_i is independent of potential citation durations $\tilde{Y}_{ij}$, 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.[1] Also, the full support condition (8) 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.

Let

$$\begin{aligned} H_i(b) = 1(Y_{i1} \leq Y_{i2})[X_{i1} - X_{i2}] \frac{\exp(X_{i2}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)} \\ + 1(Y_{i1} \geq Y_{i2})[X_{i2} - X_{i1}] \frac{\exp(X_{i1}'b)}{\exp(X_{i1}'b) + \exp(X_{i2}'b)}. \end{aligned} \tag{A3}$$

Define

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

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 $n \to \infty$,*

$$\sqrt{n}(\hat{\beta} - \beta) \to_d \mathbf{N}(0, \Omega). \tag{A4}$$

The proof of Theorem A.1 is omitted and it can be proved as in the proof of Theorem 1 of Lee (2007). The asymptotic variance Ω can be consistently estimated by

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

[1]However, this restriction can be violated if there is a cohort effect on cited patents such as technology waves.

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

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

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

and

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

B Additional Data Description and Results

In this Appendix we include several tables showing additional results.

Table A1 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 observe two citations would lead to potential selection bias. Table A2 shows the number of patents that are censored, by industry. Table A3 splits this down into the early and late period, clearly showing that the censoring problem is much more significant in the later period. Table A4 shows this by cited country. This motivates our use of estimators that explicitly allow for censoring.

In investigating the change in home bias over time we have chosen 1990 as a cutoff year because this approximately balanced the number of citations in early and later years. In Table A5 and A6 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 using also the third citation and quasi-difference between the second and third citation and we show the results from doing this in Table A7. Similarly we can use up to the fourth citation (see Table A8). Our results are robust to using these alternative citations.

Table A1: Raw data: home bias in first citation

Cited:	Citing: DE	FR	GB	EU	JP	US	RW
DE	**29.45**	2.99	2.79	4.94	13.48	40.58	5.78
FR	8.01	**18.97**	3.21	5.01	11.86	46.96	5.98
GB	7.78	3.11	**16.12**	4.48	11.44	51.13	5.93
EU	8.40	3.33	2.68	**20.68**	11.66	46.10	7.16
JP	5.32	1.76	1.59	2.73	**48.85**	34.69	5.05
US	4.42	1.93	2.06	2.84	9.18	**74.12**	5.45
RW	6.24	2.40	2.16	4.21	10.31	50.33	**24.34**

Notes: Data consists 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 2.99% of patents from German inventors were first cited by an inventor in France.

Table A2: Censoring - many patents have not (yet) been cited

	Chemicals	Computer	Drugs	Electrical	Mechanical	Other	Total
obs 2 cites	254,301	157,335	99,137	233,766	285,073	276,058	1,305,670
	(59.26)	(63.69)	(54.08)	(63.74)	(60.99)	(62.04)	(61.05)
obs 1 cite	62,495	27,764	23,346	48,917	70,421	64,587	297,530
	(14.56)	(11.24)	(12.73)	(13.34)	(15.07)	(14.52)	(13.91)
obs no cites	112,317	61,936	60,846	84,061	111,929	104,313	535,402
	(26.17)	(25.07)	(33.19)	(22.92)	(23.95)	(23.44)	(25.04)

Notes: Each row indicates the number of observations that had at least two cites ("obs 2 cites"), one cite ("obs 1 cite") or no cites ("obs no cite"). The number in parentheses indicates the proportion of observations by industry that had different numbers of cites. For example, our dataset contains 254,301 cites to patents in the chemicals technology sector that had at least two cites.

Table A3: Censoring - by early and late time period

1975-1989	Chemicals	Computer	Drugs	Electrical	Mechanical	Other	Total
obs 2 cites	174,400	71,908	51,620	141,139	191,879	186,218	817,164
	75.40	89.08	78.84	82.92	77.10	79.53	79.28
obs 1 cite	28,783	5,236	6,752	16,970	31,659	27,330	116,730
	12.44	6.49	10.31	9.97	12.72	11.67	11.32
obs no cites	28,109	3,581	7,105	12,110	25,345	20,589	96,839
	12.15	4.44	10.85	7.11	10.18	8.79	9.40
1990-1999							
obs 2 cites	79,901	85,427	47,517	92,627	93,194	89,840	488,506
	40.39	51.37	40.32	47.13	42.64	42.61	44.09
obs 1 cite	33,712	22,528	16,594	31,947	38,762	37,257	180,800
	17.04	13.55	14.08	16.26	17.74	17.67	16.32
obs no cites	84,208	58,355	53,741	71,951	86,584	83,724	438,563
	42.57	35.09	45.60	36.61	39.62	39.71	39.59

Notes: This is the same as Table A2 except we now split into early and later years.

Table A4: Censoring - by cited country

cited country:	DE	FR	GB	EU	JP	US	RW	Total
obs 2 cites	98,036	36,947	39,723	54,427	237,390	762,727	76,420	1,305,670
	57.74	56.83	61.06	54.84	60.77	63.57	51.29	61.05
obs 1 cite	26,964	10,294	9,471	15,856	53,744	157,326	23,875	297,530
	15.88	15.83	14.56	15.98	13.76	13.11	16.02	13.91
obs no cites	44,800	17,769	15,857	28,961	99,487	279,815	48,713	535,402
	26.38	27.33	24.38	29.18	25.47	23.32	32.69	25.04

Notes: This is the same as Table A2 except we now split country. DE: Germany, FR: France, GB: Great Britain, EU: other European Union (Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain and Sweden), JP: Japan, US: United States, RW: Rest of World.

Table A5: Number of Rejections of No Home Bias using Sub-Samples

Estimation Method: No Fixed Effect Estimator

Cutoff Year: 1985

Technological Category	(1) All countries Early	(2) All countries Late	(3) OECD countries Early	(4) OECD countries Late	(5) EU countries Early	(6) EU countries Late
Chemical	33	25	21	16	8	6
Computers and Communications	18	15	10	10	2	1
Drugs and Medical	24	20	17	11	7	4
Electrical and Electronic	22	21	14	11	3	0
Mechanical	28	32	20	20	6	9
Others	32	35	22	23	9	9
Total	157	148	104	91	35	29
Max. # of rejections	252	252	180	180	72	72
Percentage	0.62	0.58	0.58	0.51	0.49	0.40

Notes: 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 cut-off year. The columns under "All countries" show the number of rejections for all coefficients for country dummies (42 coefficients), those under "OECD countries" 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 those under "EU countries" 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 A6: Number of Rejections of No Home Bias using Sub-Samples

Estimation Method: Censored Fixed Effect Estimator

Cutoff Year: 1985

Technological Category	(1) All countries Early	(2) All countries Late	(3) OECD countries Early	(4) OECD countries Late	(5) EU countries Early	(6) EU countries Late
Chemical	21	10	13	5	2	1
Computers & Communications	6	6	4	4	3	1
Drugs & Medical	7	11	4	7	1	3
Electrical & Electronic	5	9	4	4	0	1
Mechanical	16	9	11	4	3	0
Others	21	16	13	9	3	2
Total	76	61	49	33	12	8
Max. # of rejections	252	252	180	180	72	72
Percentage	0.30	0.24	0.27	0.18	0.17	0.11

Notes: 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 cut-off year. The columns under "All countries" show the number of rejections for all coefficients for country dummies (42 coefficients), those under "OECD countries" 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 those under "EU countries" 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 A7: Number of Rejections of No Home Bias using Entire Sample with Second and Third Citation Spells

Technological Category	Maximum number of rejections	No FE 10%	No FE 5%	No FE 1%	FE 10%	FE 5%	FE 1%	FE+C 10%	FE+C 5%	FE+C 1 %
Chemical	42	38	33	29	20	16	7	18	15	7
Computers & Communications	42	33	26	19	10	7	6	11	8	4
Drugs & Medical	42	36	34	25	19	12	4	20	15	8
Electrical & Electronic	42	34	33	26	13	11	5	16	14	8
Mechanical	42	40	39	34	25	22	18	24	19	14
Others	42	41	37	37	20	15	8	20	16	10
Total	252	222	202	170	107	83	48	109	87	51

Notes: This is equivalent of Table 6 in the main text except we use estimates based on the second and third citation (instead of the first and second citation).

Table A8: Number of Rejections of No Home Bias using Entire Sample with Third and Fourth Citation Spells

Technological Category	Maximum number of rejections	No FE 10%	No FE 5%	No FE 1%	FE 10%	FE 5%	FE 1%	FE+C 10%	FE+C 5%	FE+C 1 %
Chemical	42	37	35	34	12	10	5	12	5	3
Computers & Communications	42	33	29	21	8	5	1	7	4	1
Drugs & Medical	42	37	28	19	10	7	2	5	4	0
Electrical & Electronic	42	29	28	25	17	10	3	14	8	6
Mechanical	42	35	34	30	14	12	6	14	10	5
Others	42	39	38	35	19	15	6	13	11	1
Total	252	210	192	164	80	59	23	65	42	16

Notes: This is equivalent of Table 6 in the main text except we use estimates based on the third and fourth citation (instead of the first and second citation).

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

817	Tsutomu Miyagawa YoungGak Kim	Measuring Organization Capital in Japan: An Empirical Assessment Using Firm-Level Data
816	Andreas Georgiadis Alan Manning	Spend It Like Beckham? Inequality and Redistribution in the UK, 1983-2004
815	Guy Michaels Xiaojia Zhi	Freedom Fries
814	Ralph Ossa	Trade Liberalization, Outsourcing, and Firm Productivity
813	Karsten Kohn Alexander C. Lembcke	Wage Distributions by Bargaining Regime: Linked Employer-Employee Data Evidence from Germany
812	Yannis M. Ioannides Henry G. Overman Esteban Rossi-Hansberg Kurt Schmidheiny	The Effect of Information and Communication Technologies on Urban Structure
811	Guy Michaels	The Division of Labor, Coordination, and the Demand for Information Processing
810	Maia Güell Jose V. Rodriguez Mora Chris Telmer	Intergenerational Mobility and the Informative Content of Surnames
809	Stephen J. Redding Daniel M. Sturm Nikolaus Wolf	History and Industry Location: Evidence from German Airports
808	Anna Lipinska	The Maastricht Convergence Criteria and Optimal Monetary Policy for the EMU Accession Countries
807	Eran Yashiv	The Beveridge Curve
806	Ghazala Azmat Alan Manning John Van Reenen	Privatization, Entry Regulation and the Decline of Labor's Share of GDP: A Cross-Country Analysis of the Network Industries
805	Henry G. Overman Patricia Rice Anthony J. Venables	Economic Linkages Across Space
804	Benjamin Aleman-Castilla	The Returns to Temporary Migration to the United States: Evidence from the Mexican Urban Employment Survey
803	Eran Yashiv	Labor Search and Matching in Macroeconomics

802	Nicholas Oulton	Jeremy Greenwood and Per Krusell, "Growth Accounting with Investment-Specific Technological Progress: A Discussion of Two Approaches" A Rejoinder
801	Mirabelle Muûls Mauro Pisu	Imports and Exports at the Level of the Firm: Evidence from Belgium
800	Richard E. Baldwin Frédéric Robert-Nicoud	Protection for Sale Made Easy
799	Alejandro Cuñat Marc J. Melitz	Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage
798	Giulia Faggio	Job Destruction, Job Creation and Unemployment in Transition Countries: What Can We Learn?
797	Nicholas Oulton	Chain Indices of the Cost of Living and the Path-Dependence Problem: an Empirical Solution
796	David Marsden Richard Belfield Salima Benhamou	Inventive Pay Systems and the Management of Human Resources in France and Great Britain
795	Andrew B. Bernard J. Bradford Jensen Stephen Redding Peter K. Schott	Firms in International Trade
794	Richard E. Baldwin Frédéric Robert-Nicoud	Offshoring: General Equilibrium Effects on Wages, Production and Trade
793	Alan Manning	Respect
792	Nick Bloom	Uncertainty and the Dynamics of R&D
791	Richard E. Baldwin Frédéric Robert-Nicoud	Entry and Asymmetric Lobbying: Why Governments Pick Losers
790	Alan Manning Sanchari Roy	Culture Clash or Culture Club? The Identity and Attitudes of Immigrants in Britain
789	Giorgio Gobbi Roberta Zizza	Does the Underground Economy Hold Back Financial Deepening? Evidence from the Italian Credit Market
788	Nick Bloom Raffaella Sadun John Van Reenen	Americans do I.T. better: US Multinationals and the Productivity Miracle

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk