How Special is the Special Relationship? Using the impact of US R&D spillovers on UK firms as a test of Technology Sourcing*

Rachel Griffith†, Rupert Harrison†‡ and John Van Reenen§

June 22, 2004

Abstract

How much does US-based R&D benefit other countries? We test the hypothesis that foreign firms locating inventors in the US benefit from "technology sourcing" (i.e. that foreign research labs on US soil tap into US R&D spillovers). Using panels of British and US firms matched to patent data we show that British firms who had established a high proprtion of US-based inventors by 1990 disproportionately benefitted from the growth in the US R&D stock in the next 10 years. We estimate that UK manufacturing TFP would have been about 5% lower in 2000 (about $14bn) in the absence of the US R&D growth in the 1990s. By contrast, the growth of the UK R&D stock did not appear to have a major benefit for US firms. The "special relationship" between the UK and the US is distinctly asymmetric.

JEL No. O32, O33, F23

Keywords: knowledge spillovers; technology sourcing; productivity

*__Acknowledgement:__ The authors would like to thank Nick Bloom, Steve Bond, Jan Boone, Michele Cincera, John Sutton as well as participants at seminars at LSE, Swansea and Tilburg for helpful comments. Financial support for this project was provided by the ESRC Centre for the Microeconomic Analysis of Fiscal Policy at the IFS and the ESRC Centre for Economic Performance. The data was developed with funding from the Leverhulme Trust.

†Institute for Fiscal Studies and University College London

‡Corresponding author (rupert.harrison@ifs.org.uk)

§Centre for Economic Performance, London School of Economics and CEPR

1. Introduction

There is a consensus among economists that a substantial part of global economic growth arises from the transfer of ideas from the leading edge countries to those behind the frontier. But the mechanisms underlying this technology transfer are poorly understood and micro-econometric evidence on the quantitative importance of the "international spillover" process remains thin.[1] In addition, the firm level evidence on spillovers that does exist tends to be from single countries and the bulk of these single country studies are from the United States, which, as technological leader in most industries, probably has least to gain from other countries' innovative efforts.

Case studies and the business press have long emphasized the importance of "technology sourcing" as a method of gaining access to foreign knowledge[2]. Under this view, firms can tap into leading edge knowledge by setting up R&D labs in the US which act as "listening posts" for new ideas. The main contribution of our paper is to provide the first rigorous evidence for technology sourcing by exploiting firm level panel data from the UK and the US. UK firms offer a particularly good testing ground for these theories because Britain is both less technologically advanced than the US[3] and has historically close linkages to US based inventors[4].

[1] Ses Wolfgang Keller (2004) for a recent survey.

[2] See for example von Zedtwitz and Gassman (2002) or Serapio and Dalton (1999) and the references therein.

[3] In the market sector (i.e. excluding health, education and public administration) output per person was about 40% higher in the US than in the UK in 1999 (US TFP was about 20% higher).

[4] Of all foreign countries, British expenditure on R&D in the US was second in the world only to Switzerland in 1993. In 1997, of the largest 7 foreign research centres in the US, five

We examine whether the US R&D stock (conditional on British R&D) had a stronger impact on the TFP of UK firms who had more of their inventors located in the US. We use the pre-1990 location patterns of British firms (as revealed in the patent statistics) to mitigate the endogeneity problem arising from the fact that British firms may choose to locate R&D in the US in response to the 1990s technology boom.

We illustrate our identification strategy in Figure 1. The horizontal axis shows the growth of the US R&D stock by industry between 1990 and 2000. On the vertical axis we plot the "productivity premium" for UK firms who had a substantial proportion of US inventors (i.e. the difference in productivity growth for UK firms with a high proportion of their inventors located in the US prior to 1990 vs. UK firms with zero or low US inventor presence). It is clear that the productivity premium is larger in those industries where the US had faster R&D growth. Furthermore, the shaded industries are those where the US already had a substantial technological lead over the UK by 1990 and where, presumably, UK firms had the most to learn. For this "high gap" sector, the upward sloping relationship is particularly striking.

[Figure 1 about here]

The graph does not control for many other confounding influences and the paper uses a variety of econometric methods to deal with input endogeneity, unobserved heterogeneity and selectivity. Even after controlling for these issues, we

were owned by UK companies (Serapio and Dalton, 1999). In our data almost half of patents granted to UK firms were produced by inventors located in the US.

find that UK firms which had more of their inventive activity located in the US *prior* to 1990, benefited disproportionately from the burst in US R&D growth in the 1990s. In fact we find *no* significant impact of US R&D on British firms who have no US-based inventors. According to our estimates, TFP in British manufacturing in 2000 would have been 5% lower (about \$14bn)[5] in the absence of the growth of US R&D stock in the 1990s. Needless to say, this is a lower bound to the full benefits of US R&D to the rest of the world. It is also a salutary warning to policy makers who seek to boost sluggish European growth through incentivising multinationals to repatriate US R&D back towards Europe[6].

Our research has links to several strands in the literature. First, there is much work suggesting that knowledge spillovers are partly localised and that being geographically close to innovators matters.[7] We build on this work by focusing on the location of inventors *within* firms across geographic boundaries. Second, except for some aggregate studies[8] the work on multinationals focuses on the benefits to the *recipient* country of inward FDI.[9] By contrast, we examine whether outward innovative FDI to specific industries in a leading edge country have beneficial affects on home country productivity. Thirdly, although there

[5] Value added in UK manufacturing was £154bn in 2000=\$277bn at current exchange rates

[6] The European Union has set itself the target of increasing R&D expenditure to 3% of GDP by 2010 (this is part of the "Lisbon Agenda").

[7] For example, Adam Jaffe et al (1993, 2000), Wolfgang Keller (2002), David Audretsch and Marion Feldman (1996). Adam Jaffe and Manuel Trajtenberg (1998) find that, even after controlling for other factors, inventors residing in the same country are typically more likely to cite each other than inventors from other countries, and that these citations tend to come sooner. They also find that localisation fades over time, but only slowly.

[8] For example, Frank Lichtenberg and Bruno van Pottelsberghe de la Potterie (2001)

[9] For example, see Wolfgang Keller and Yeaple (2003) for recent US evidence.

is some recent research that has examined the evidence for technology sourcing through patent citations,[10] we are aware of *no* studies that consider empirical evidence for technology sourcing in terms of its effects on firm-level productivity.[11] We also show that cross country patent citations (at the firm level) are consistent with our results, but we believe that the impact of US technology on foreign firm performance may not be fully revealed in patent citations as some of the knowledge created is tacit rather than codified - this is captured in our TFP results.

We contrast our UK production functions with identical specifications based on US firm level panel data. Although it is possible that US firms technology source from the UK, it is much less likely to be important, as the UK is generally not at the technology frontier. This is indeed what we find. The "special relationship" between the UK and the US is asymmetric: Britain benefits more, at least in respect of knowledge flows.

The structure of this paper is as follows. Section 2 sketches the empirical model and Section 3 describes the data. Section 4 presents the empirical results,

[10] Branstetter (2003) uses patent citations to measure the role of foreign direct investment by Japanese firms in the USA in mediating flows of knowledge between the two countries. He finds that knowledge spillovers received by the investing Japanese firms tend to be strongest via R&D and product development facilities which is consistent with our findings. Iwasi and Odagiri (2002) claim that Japanese research facilities foster the innovative activity of the investing parent firm, but they only have cross sectional evidence. Singh (2003) uses patent citations to investigate the role of multinational subsidiaries in knowledge diffusion. He finds that greater multinational subsidiary activity increases cross-border knowledge flows between the host country and the multinational home base.

[11] Lee Bransetter (2001) enters the US R&D pool in a Japanese production function and finds a positive, but insignificant coefficient. He does not allow the effect to differ with Japanese inventor presence in the US, however (a test of technology sourcing). In addition, the author is not confident in the quality of the Japanese R&D stock data, because of the short time span (p.72).

and a final section concludes. The details of the data and econometric methods are in the Annexes.

2. The empirical model

Our basic approach follows Zvi Griliches (1979) and many subsequent papers by including measures of the external knowledge stock available to the firm in a firm-level production function. A firm's value-added is a function of traditional inputs as well as knowledge and can be written as follows

$$Y = F(X, DOMESTIC, FOREIGN) \tag{2.1}$$

where Y is real value added, X is a vector of the firm's own inputs including labour, capital and the firm's own knowledge stock accumulated by performing R&D. $DOMESTIC$ is the domestic external knowledge stock available to the firm and $FOREIGN$ is the foreign external knowledge stock.

The key test in this paper is whether the effect of the external knowledge stock on productivity depends on the geographical location of the firm's innovative activity (denoted W). In particular we are interested in spillovers from the foreign external knowledge stock

$$\frac{\partial Y}{\partial FOREIGN} = f(W^F) > 0; \tag{2.2}$$

and we want to test whether this is increasing in the size of a firms presence in the foreign location, i.e.

$$\frac{\partial^2 Y}{(\partial FOREIGN)(\partial W^F)} > 0. \tag{2.3}$$

where W^F measures of the amount of the firm's innovative activity that is located abroad. Analogously, we allow the domestic R&D spillovers to be more important if the firm's inventive activity is mainly located domestically (W^D).

To make matters more concrete consider a Cobb-Douglas production function for firms in the UK with the US being the relevant foreign source of spillovers

$$Y_{it} = A_{it} L_{it}^{\alpha_l} K_{it}^{\alpha_k} R_{it}^{\beta} DOMESTIC_{jt}^{\gamma_{i1}} FOREIGN_{jt}^{\gamma_{i2}} \tag{2.4}$$

where i indexes a firm, j indexes the firm's industry, and t indexes the year. Y_{it} is real value added, L_{it} is employment, K_{it} is a measure of the firm's capital stock, R_{it} is a measure of the firm's own R&D stock, and $DOMESTIC_{jt}$ and $FOREIGN_{jt}$ are the R&D stocks in the firm's industry in the UK and the USA respectively.[12] We further assume that the elasticities of value added with respect to the domestic and external knowledge stocks are a linear function of firm-specific measures of the location of innovative activity[13]

$$\gamma_{i1} = \theta_1 + \theta_2 W_i^{UK}; \gamma_{i2} = \phi_1 + \phi_2 W_i^{US}; \tag{2.5}$$

where we interpret a positive estimate of ϕ_2 as evidence of knowledge spillovers associated with technology sourcing from the USA. Using lower case letters ($x = lnX$, etc.) to denote natural logarithms we obtain:

[12] We looked at also including other countries R&D stocks (in addition to the US) and their interactions in $FOREIGN$, but although usually positive these were not significantly different from zero. This is not to say that the UK learns only from the US, rather that the US is by a long way the most important partner. We also consider the measure of spillovers using in Jaffe (1984) and our results were robust to this.

[13] We consider more flexible functional forms in the results section, but these do not change the main qualitative results.

$$
\begin{aligned}
y_{it} &= \alpha_l l_{it} + \alpha_k k_{it} + \beta r_{it} + \phi_1 foreign_{jt} + \phi_2 (W_i^{US} * foreign_{jt}) \\
&\quad + \theta_1 domestic_{jt} + \theta_2 (W_i^{UK} * domestic_{jt}) \\
&\quad + \phi_3 W_i^{US} + \theta_3 W_i^{UK} + a_{it} \qquad (2.6)
\end{aligned}
$$

Our baseline equations are for the UK, but we estimate an identically specified equation for the US to see if technology sourcing from the UK also matters for US firms. We expect it to matter a lot less as the UK is not at the technological frontier.

There are a number of econometric issues involved in estimating firm level production functions such as (2.6), the basic issue being how to deal with the endogeneity of the firm's choice variables in the presence of unobserved heterogeneity. Our basic approach follows the "System" General Method of Moments (SYS-GMM) approach of Richard Blundell and Stephen Bond (2000). We also compare these results with those from OLS and an extension to the Olley-Pakes (1996) method which allows for endogenous R&D. Econometric details are contained in Appendix B, but we note some features here.

The generic problem of estimating a firm production function is that the firm's inputs choices are likely to be correlated with the productivity shock, a_{it}. Under SYS-GMM we assume that the residual terms can be broken down into $a_{it} = t_t + \eta_i + u_{it}$ where year dummies (t_t) control for common macro effects, the firm fixed effect (η_i) controls for unobserved heterogeneity and the residual productivity shock (u_{it}) may be correlated with the regressors. Assumptions over the initial conditions yield moment conditions for the levels equations which can be combined in a system using the traditional moment conditions for the first differenced equations (generated by assumptions over the serial correlation properties of the u_{it} term). In both equations we essentially use lagged values to construct instrumental variables for current variables.

The Olley Pakes (OP) algorithm is based on a structural model which generates a two step method. In the first step we obtain a consistent estimate of the labour coefficient (α_l) using a non-parametric approach to sweep out with the correlation of variable inputs with the error term. In the second step we obtain the capital parameter using non-linear least squares. The routine avoids using instrumental variables, but does not extend so straightforwardly to endogenous R&D decisions. We therefore consider an extension to Olley-Pakes which allows for endogenous R&D following Thomas Buettner (2004). This leaves stage one intact of the algorithm, but alters the way we draw inferences on the capital coefficient at stage 2.

Whether we use OLS, GMM or OP we still have the intrinsic problem that the coefficients on our R&D spillover terms may reflect other correlated shocks to demand or supply.[14] We attempt to control for such biases by including firm (or industry) fixed effects and other industry variables (such as sector-level demand terms). We also examine using lags of the spillover terms, which should be less effected by contemporaneous shocks. Of course the key variable for us is the

[14] See Charles Manski (1991) for a general discussion. Note that this is more likely to be a problem for the coefficients on the domestic R&D spillover terms (θ_1, θ_2) than the foreign R&D spillover terms since UK firms mainly produce domestically.

coefficient on the interaction term between foreign R&D and the location weight (ϕ_2). There is no obvious reason why there would be an upwards bias to this term, even if there was upwards bias to the linear spillover term (ϕ_1).

A related concern is that W_i^D and W_i^F are choice variables for the firm (or at least reflect past R&D locational choices), and may thus be correlated with firm or industry-level technological shocks in a way that undermines our identification strategy. Since we have no convincing exogenous instruments for the location of firms' innovative activity we use *pre-sample* information to construct W_i^D and W_i^F. This ensures that the locational variables are not affected by shocks that also directly affect firm-level outcomes during the sample period.[15] This strategy assumes that the firm did not locate R&D in the US in anticipation of positive shocks to productivity. While we cannot rule out such behaviour, the fact that the inventor patents are the result of R&D decisions taken often decades prior to the sample period makes such biases likely to be small.

A final worry is that our empirical measure of W_i^F may be proxying for other non-locational aspects of firm's activities (e.g. absorptive capacity or unobserved firm quality) or non-innovation related aspects of the firm (e.g. its US production activities). We test carefully for these alternative explanations in the results section by bringing other types of data to bear upon the problem, such as citations information.

[15] This has the disadvantage that firms may have moved their inventive activity over time. This should, however, bias aganist us finding evidence of technology sourcing.

3. Data

Our main dataset is a panel of manufacturing companies listed on the London Stock Exchange. We matched information on all the US patents taken out by these firms since 1968 (using the NBER/Case Western Patents dataset) by name matching. These firms account for approximately 80% of all UK R&D in 1999. Table 1 shows that firms in our sample had 63,733 US patents (which made 472,998 citations). 31% of these patents had inventors located in the UK compared to only 3% in the USPTO population as a whole - this is unsurprising. since these are all firms listed on the London Stock Exchange. A further 45% of the patents taken out by our UK firms had inventors located in the US - this illustrates the importance of the US as a home for inventive activity of UK firms. But it also reflects the fact that we are using US patents rather than UK or European patents. Our US firm data is based on the match between Compustat and the USPTO conducted by Bronwyn Hall et al (2001). The distribution of inventors in these firms is contained in the third column and shows that only 1% of inventors were located in the UK compared to 92% in the US itself. This illustrates why it would be hard to examine technology sourcing from US data alone.

Table 2 gives some further descriptive statistics on our UK sample. Since all these firms perform R&D and are listed on the Stock Exchange they are larger than typical UK firms (the median employment is 1,750). Compared to the sample of US firms (see Table A5), however, the UK firms are smaller (median US firm size is 3,528). UK firms are also less R&D intensive that their US counterparts which mirrors the aggregate statistics. The full details of the data construction are in Appendix A.

We use several measures of W_i^{UK} and W_i^{US}. The basic measure is constructed as the proportion of the firm's total patents applied for between 1975 and 1989 where the inventor is located in the UK or the USA respectively. They are both equal to zero if the firm has no patents. Our firm panel runs from 1990 to 2000, so the location measures are based purely on *pre-sample information.* As discussed above, this ensures that the location measures are not affected by shocks that affect firm-level outcomes during the sample period.

This measure of the geographical location of innovative activity discards two types of information in the patent data. The first is variation over time, so that the measure represents an average of the location of the firm's innovative activity over the period 1975-1989.[16] The second type of information is the total number of the firm's patents. While this may be relevant information, normalising the location measures to a proportion between zero and one helps to deal with difficulties associated with firm size and differences in propensity to patent across industries. In the results section we show that the results are robust to adding in extra terms of the interaction of the spillover terms with functions of the number of patents.

As mentioned above, we also use information on patent citations to refine our measure of W_i^{UK} and W_i^{US}. A key theme in the literature is that technology sourcing is not the only motivation for firms to locate innovative activity abroad. In particular, firms may do R&D abroad in order to adapt existing technologies

[16] We also constructed W_i using data only in the 1990s which gave similar results.

to new markets. Our empirical approach to this issue is to use data on citations to eliminate patents that are unlikely to represent technology sourcing behaviour. Consider two extreme cases for a patent that is owned by a UK firm but that was invented in the US: if the patent only cites patents owned by the same parent firm and whose inventors were located in the UK then the patent is more likely to represent activity associated with adapting an existing technology to the US market. On the other hand, if the patent cites many patents that are not owned by the parent firm and whose inventors were located in the US then the patent is more likely to represent technology sourcing behaviour. If we want to investigate whether there is evidence for technology sourcing behaviour in productivity outcomes, then we want to focus on the latter and not use the first type of patent when constructing our location measures.

To implement this approach, our second measure of W_i^{UK} and W_i^{US} looks only at patents that cite patents whose inventors were located in the same country and were not owned within the same parent firm. This measure of W_i^{US} is equal to the proportion of the firm's total patents where: (1) the inventor is located in the USA and (2) the patent cites at least one other patent whose inventor was both located in the US and did not work for the same parent firm.

Our third and most refined measure of W_i^{UK} and W_i^{US} is the same as the second measure, except that it also uses information on the time-lag between the citing and cited patent. Technology sourcing behaviour is likely to be associated with gaining access to pools of "tacit" knowledge. Given that knowledge that was created recently is more likely to have tacit characteristics, we include only citations to patents whose application date is no more than three years prior to that of the citing patent. The third measure of W_i^{US} is thus equal to the proportion of the firm's total patents where: (1) the inventor is located in the USA and (2) the patent cites at least one other patent that was applied for within the last three years and whose inventor was both located in the US and did not work for the same parent firm.

4. Results

4.1. Production Function: Main Results

Table 3 presents our main results on our R&D augmented production functions. Columns (1) and (2) present the OLS results; column (1) does not impose constant returns to scale in labour and capital, while column (2) does.[17] Columns (3) through (5) present System-GMM results. Column (3) contains the basic measure of location (e.g. the proportion of inventors based in the US) whereas the next two columns present the closer refinements to technology sourcing based on citation patterns. In all columns the coefficient on the labour-capital ratio is similar to the OLS case (about 0.65, close to labour's share in value added). The estimated elasticity with respect to firm-specific R&D is positive and corresponds to an private excess rate of return to R&D of about 15% for our average firm, which is similar to that found in other studies.[18] Diagnostic tests are presented for first and second order serial correlation in the first-differenced residuals. Neither test ever

[17] CRS is not rejected in the SYS-GMM results and is marginally rejected for OLS.

[18] For example, Griliches (1992). The private rate of return is calculated as $\widehat{\beta} * (\frac{Y}{R})$ which at the average UK firm's R&D stock intensity is 0.025*6 = 0.15

rejects the hypothesis of no serial correlation. This justifies the use of twice lagged instruments in the difference equation and once lagged instruments in the levels equation.[19] A Sargan test of the overidentifying restrictions is not significant at 5%, and neither is a Sargan difference test of the extra moment conditions implied by the levels equation.

Turning to our main variables of interest, the key interaction term (ϕ_2) between US inventor location (W_i^{US}) and the US R&D stock ($Foreign_{jt}$) is positive and significant at conventional levels across all the UK specifications. This seems consistent with our technology sourcing argument. The linear US R&D term is positive, but insignificant which implies that a UK firm with no inventor activity in the US would receive *no* spillovers from US R&D.[20] The basic domestic UK industry R&D term enters positively and significantly, but the interaction with W_i^{UK} is not significantly different from zero. This may imply that UK firms do not have to have a significant number of inventors located in the UK in order to take advantage of domestic R&D spillovers.

Column (4) refines the location weight by only including patents which were not self-citations and which did cite at least one other US inventors, as discussed in the previous section. Column (5) further refines the measure by including only patents that have cited other inventors in the last three years. The two refinements bring the measure of inventor location closer to the theoretical ideal of technology

[19] In addition, none of the key results are sensitive to dropping twice-lagged differences and once-lagged levels from the instrument set.

[20] This is probably too strong as some spillovers are likely to be captured by the time dummies and also transferred through non-R&D channels.

sourcing, although at the cost of using thinner slices of the patents data (see table A3). It is reassuring that the coefficient on our key interaction ($W_i^{US} * Foreign_{jt}$) becomes increasingly strong as we move from column (3) through (5). This is consistent with the notion that the measures are capturing what we intend rather than some other spurious relationship.[21]

Column (6) reports the Olley-Pakes estimates of the production function. The coefficient on labour is a bit lower relative to the OLS and the coefficient on capital is a bit higher.[22] The OLS bias is what one would expect from endogeneity of inputs and selectivity.[23] The qualitative findings are robust, however, and the interaction between US R&D and foreign location remaining highly significant.[24]

Overall, there appears to be strong evidence that the productivity growth of UK firms is significantly higher in industries with strong US R&D growth if and only if the UK firms already have an inventive presence in the US. This is

[21] It is interesting that the linear US location measures W_i^{US} are usually negative suggesting that there is some costs to locating inventors outside the home country (although note that this term enters positively when the interactions are not included). The median marginal effect of W_i^{US} on value added remains positive (e.g. in column (3) the median marginal effect is 0.03, and the median marginal effect is positive in 10 out of 15 industries).

[22] The coefficient on firm R&D - although positive - is quite a bit lower than the other estimates. This may be because the methodology already controls for R&D at stage 1 (see Appendix B).

[23] Endogeneity of input choice generally leads to an upward bias on the labour coefficient and a downward bias on the capital coefficient as there is generally a higher contemporaneous correlation between labour and productivity than between capital and productivity (Marschak and Andrews, 1944; James Levinshohn and Amil Petrin, 2003).

[24] The OP results are generated by a multi-stage procedure (see Appendix B for details). We first estimate the coefficient on labour using a control function approach. Using this (and a selection equation) we then estimate the coefficient on capital by non-linear least squares. Given the estimates of these we then calculate firm level efficiency and relate this to the spillover terms as shown. An alternative method with introduces the spillovers terms directly into stage 2 (analgously to "plant age" in Olley and Pakes, 1996) gave qualitively similar but less precise results.

consistent with the technology sourcing hypothesis.

4.2. Production Function: Further Investigations

4.2.1. US Results

All these results up until now are for UK firms. Given that the US is usually at the technology frontier and UK firms are generally below the technology frontier we might expect that technology sourcing is a much less powerful motivation for US firms locating their R&D labs in the UK. Indeed, the proportion of US firms' patents that have a UK inventor is only about 1% (compared to 45% of UK firms patents with a US inventor). This is why using a sample of UK firms is a much better test of the technology sourcing hypothesis then using US firms. In order to investigate this, column (1) of table 4 re-estimates the specification in column (3) of Table 3 for US firms (see Appendix A for details of the data). The coefficients on labour and capital are similar to the GMM estimates for the UK. The domestic US R&D term is positive and significant suggesting domestic spillovers, but the interaction with the location weight is insignificantly different from zero. Consistent with our expectations, the interaction between UK presence and UK R&D ($W_i^{UK} * Domestic_{jt}$) is insignificantly different from zero (although it is positive).[25] Even if the interaction where statistically significant, however, the economic magnitude of the impact is small. A US firm would have to have at least half of its inventors in the UK before UK R&D achieved any positive

[25] These results are robust to dealing with spillovers in many different ways. For example, constructing a Jaffe (1984) measure of spillovers for the US firms and interacting it with domestic inventor presence was insignificant.

productivity impact (only 0.5% of the US sample are in this position).

4.2.2. Location of Productive Activity

Although we have been assuming that production activity is located in the UK, this is not completely true in practice. It is possible that the location measure W_i^{US} is not only proxying for the location of *innovative activity*, but also for the location of *production*. In other words, firms with innovative activity in the USA may also have productive activity located there. If this is the case, then we may be picking up not only international spillovers but also domestic spillovers within the USA. We attempt to control for this by estimating our model on firms with no (or practically no) US production activities (72% of our firms are in this category) based on their reported number of domestically and overseas employees.[26] Looking only at firms whose productive activity is located entirely within the UK the results are very similar - the key interaction of inventor location with US R&D stock has a coefficient of 0.221 and standard error of 0.063 (see column (2) of Table 4), actually slightly stronger than in Table 3.[27] This suggests that our UK results are not primarily driven by the location of firms' production activities.

4.2.3. Industry Heterogeneity

[26] 117 out of 188 firms report domestic employment separately to total employment at least once during 1990-2000. For those that do not report separately we assume that all employment is domestic. Of those 117 firms, 53 report total employment greater than domestic employment at least once. We drop these firms from the sample and re-estimate our model on the remaining 135 firms, which we expect to have little or no foreign production activity.

[27] This is from a specification identical to that of Table 3 column (5).

We examined whether the technology sourcing effect was different across industries. Since the theory suggests that technology sourcing should be stronger for UK firms which are furthest behind the technological frontier we divided up industries into those where the TFP gap with the US was large versus those where the TFP gap was smaller. We found that the US interaction term was much stronger in the sectors where the UK firms "had the most to learn" from the US. This is illustrated in columns (3) and (4) of Table 4: the key coefficient is twice as large and only significant in the "high TFP gap industries".

4.2.4. Absorptive Capacity

One interpretational difficulty arises if the inventor location term reflects the firm's own intensity of knowledge. For example, if firms located in the US are more innovative and if innovative firms absorb knowledge more easily, this could account for the positive interaction. To test this we included further interactions of the spillover measures with indicators of the firms overall innovativeness. Although these were generally positive they were less informative than the location interactions. For example, we interacted a binary dummy indicating whether the firm had ever patented with the industry R&D terms. This is to check that the results on the location interactions are not driven by patenting firms having higher "absorptive capacity" than non-patenting firms, since non-patenting firms by definition have values of W_i^{UK} and W_i^{US} equal to zero. Neither of the interactions with the patenting dummy is ever significant, and the positive significant interaction with W_i^{US} remains, suggesting that the results are not driven by absorptive capacity.

The concern over absorptive capacity is similar to the concern that the W_i^{US} reflects some other form of unobserved heterogeneity[28]. We conduct a further test of this using a patents citation equation in the next section.

4.2.5. Other Robustness Tests

We also conducted a large number of other robustness checks. First, we included industry level value added in the US and in the UK to check that the results are not driven by industry level shocks correlated with R&D. None of the value added terms is significant in the UK equations.[29] Second, we lagged all the industry level R&D terms by one period, so that they could be considered pre-determined. Again the main results are not affected. Thirdly, there is a worry that we have omitted the human capital composition of the firm as such data is unavailable from company accounts. To check for this we included the average wage and its interaction with US R&D to make sure this was not driving the results. Our key interaction remained significant and positive (although the linear average wage term was significant).

4.3. Patent Citation Equations

Our interpretation of the results in the previous section is that having inventors located in the US allows UK firms to access geographically localised spillovers.

[28] We also interacted a pre-sample measure of firm specific TFP to see if this was driving the results. It was insignificant.

[29] US value added was significant in the US firms production function and we keep it in the Table 4 column (1) results. We also included interactions of value added with the locational weights - none of these interactions were significant.

However, it is possible that the firm-level location weights are correlated with some unobserved firm-level characteristic that allows firms to absorb the information contained in spillovers from the US. For example, some UK firms may be located in a technological area in which US firms have a comparative advantage. Recently, many authors have turned to patent citations as an alternative, direct way of measuring spillovers.[30] We use this alternative source of information as one way of investigating the possibility that our previous results are driven by unobserved heterogeneity.

To implement this approach we estimate a patent citation equation of the following form.

$$CITES_{pit}^{US} = g(US_{pit}, UK_{pit}, W_i^{US}, W_i^{UK}, x_{pit}, u_{pit}) \tag{4.1}$$

The dependent variable $CITES_{pit}^{US}$ is a count of the number of non-self citations from patent p of UK firm i at time t to a patent with a US inventor that was applied for within the previous three years. This is the type of citation that we consider most likely to be associated with technology sourcing. US_{pit} and UK_{pit} are dummy variables that are equal to unity if the citing patent is invented in the US or UK respectively, and zero otherwise. The base category is all other countries. W_i^{US} and W_i^{UK} are the basic firm-level location weights described above. Control variables (x_{pit}) include the total number of cites made by the patent ($TOTALCITES$), year dummies, industry dummies, technology class dummies and all other firm and industry-level variables in the production

[30] For an early example see Adam Jaffe, Manuel Trajtenberg and Rebecca Henderson (1993).

function. Finally, u_{pit} is a serially uncorrelated error term.

An established result in the citations literature is that patents are more likely to cite other inventors in the same country than they are to cite foreign inventors, and these citations tend to come sooner.[31] Thus we expect US_{pit} to enter positively in equation (4.1). However, if our interpretation of the production function results is correct, we expect the firm-level variable W_i^{US} not to enter in equation (4.1) conditional on the location of the citing patent's inventor. If W_i^{US} were to enter positively, even after controlling for US_{pit}, this would suggest the presence of some firm-level propensity to cite US inventors that was not entirely accounted for by the presence of individual inventors in the US. In particular, it might be the case that the firm's UK-based inventors were also systematically more likely to cite US inventors. This would suggest that the firm-level location weight W_i^{US} was proxying for something more than just the geographical location of inventors.

The sample is all patents applied for by our sample of UK firms over 1990-1998. Restricting our attention to patents applied for after 1989 allows us to use the same pre-sample firm-level location weights as before.[32] We estimate equation (4.1) by a negative binomial count data model, and as a robustness check we also estimate a probit regression where the dependent variable is equal to one if $CITES_{pit}^{US}$ is greater than zero.[33]

[31] See Adam Jaffe and Manuel Trajtenberg (2002) for a recent survey of this literature.

[32] We do not consider patents applied for after 1998 because the patent database only contains information on granted patents. Since the process of granting a patent can take several years, this raises the possibility of truncation bias by not including patents that have been applied for but not yet granted.

[33] Similar results to the Negative Binomial model emerge from a Poisson specification, although the Poisson model is strongly rejected in favour of over-dispersion. The data support a hypothesis

Table 5 presents the results. In column (1) we exclude the individual inventor location indicators US_{pit} and UK_{pit}. The firm-level location variable W_i^{US} is strongly associated with the propensity to cite US inventors. This initial result is reassuring as it corroborates the hypothesis that our firm-level inventor location weight is picking up knowledge transfers using a completely different methodology to the production function approach. If the US R&D labs of our UK firms were not really tapping into localised US knowledge (e.g. if they were just adapting European knowledge to the US market) we would not expect them to be extensively citing US patents.

In column (2) we include US_{pit} and UK_{pit} in the specification. The coefficient on the US inventor dummy is positive and highly significant, confirming the result found elsewhere in the literature that US inventors are more likely than foreign inventors to cite other US inventors. This is true even though all the patents in the sample are owned by UK firms. The reported coefficients on US_{pit} suggests that the citation rate per patent to US inventors is about 68% higher for patents invented in the US. More importantly for our purposes, conditioning on the location of the patent's inventor drives the coefficient on the firm-level location weight W_i^{US} to zero. So there is no evidence for any firm-level propensity to cite US inventors that is not entirely accounted for by the presence of individual inventors in the US. In particular, UK inventors are not more likely to cite US inventors when their firm's value of W_i^{US} is high. Columns (3) and (4) present equivalent

of constant dispersion, with the additional dispersion coefficient, delta, significantly greater than zero, as shown in Table 4.

results for the probit specification. As with the negative binomial results the introduction of US_{pit} drives the coefficient on W_i^{US} towards zero.

These results from patent citation behaviour tend to support our interpretation of the earlier production function results. UK firms with inventors located in the US are more able to benefit from localised US spillovers precisely because of the presence of those inventors in the US, and not because of some other firm-level characteristic that is correlated with having inventors located in the US.

5. Summary and Conclusions

The results presented in this paper provide strong evidence for the existence of knowledge spillovers associated with technology sourcing. The idea that firms might invest in R&D activity in a technologically advanced country such as the US in order to gain access to spillovers of new "tacit" knowledge has been suggested in the business literature but we know of no studies that have attempted to find evidence for this in observed productivity outcomes.

Our main results suggest that the increase in the US R&D stock in manufacturing over 1990-2000 was associated with on average a 5% higher level of TFP for the UK firms in our sample (about $77bn at 2000 prices). This compares with an average 6% higher level of TFP associated with the increase in their own R&D stocks over the same period.[34] Thus spillovers from the US contributed about

[34] These numbers are calculated as the product of the estimated elasticities from Table 3 and the percentage change in the US and own R&D stocks over the 1990-2000 period. All three location weights gave similar estimates of the contribution of US R&D to the TFP growth of our sample of firms.

two-thirds of the effect of firms' own R&D. Our results also suggest that for a UK firm, shifting 10% of its innovative activity (as measured by patent applications) to the US from the UK while keeping its overall level of R&D stock the same (e.g. changing W_i^{US} from 0.30 to 0.40 and W_i^{UK} from 0.70 to 0.60 while keeping R_{it} the same) is associated with an increase in its TFP level of about 3%. This effect is the same order of magnitude as that of a doubling in its R&D stock.

The US innovation boom in 1990s had major benefits for the UK economy (and by implication for many other countries in the world). An interesting extension of our methods would be to replicate the findings from other countries. It also increases the incentives for multinationals to locate R&D in the US, which is indeed what has occurred. Future research needs to show to what extent this is driven by technology sourcing rather than other contemporaneous events (such as the increasing generosity of the US R&D tax credit).

Our result has interesting implications for policy. Governments are generally keen to promote higher levels of domestic R&D activity, and the Member States of the European Union have recently expressed an aspiration to raise the level of R&D spending within the EU to 3% of GDP. Policies which aggressively seek to achieve this target by artificially inducing multinational European firms to relocate their existing R&D labs away from the US and towards Europe could be very counterproductive as they may reduce the ability of European firms to benefit from US R&D spillovers.

6. References

Audretsch, David and Feldman, Marion (1996) "R&D Spillovers and the Geography of Innovation and Production" *American Economic Review*, 86(4), 253-273

Bloom, Nick and Van Reenen, John (2002), "Patents, Real Options and Firm Performance", *The Economic Journal* 112, 478, C97-C116

Blundell, Richard and Bond, Stephen (2000), "GMM estimation with persistent panel data: An application to production functions", *Econometric Reviews* 19(3), 321-340

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

Branstetter, Lee (2003), "Is foreign direct investment a channel of knowledge spillovers? Evidence from Japan's FDI in the United States", Columbia Business School and NBER, mimeo

Buettner, Thomas (2004), "R&D and the Dynamics of Productivity" London School of Economic mimeo

Griliches, Zvi (1979), "Issues in assessing the contribution of research and development to productivity growth", *Bell Journal of Economics* 10 (1), pp 92-1167

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

Griliches, Zvi (1994), "Productivity, R&D, and the data constraint", *American Economic Review* 84, (1), pp 1-23

Griliches, Zvi (1996), "R&D and productivity: The unfinished business", in Z.Griliches (1998), R&D and Productivity: The Econometric Evidence, The University of Chicago Press, Chicago and London

Hall, Bronwyn, Jaffe, Adam and Manuel Trajtenberg (2001) "Market Value and Patent Citations: A first look" Univesity of California, Berkeley Working Paper No. E01-304

Iwasi, Tomoko and Hiroyuki Odigari (2000) "The role of overseas R&D activities in technological knowledge sourcing: An Empirical study of Japanese R&D investment in the US" Discussion Paper No. 23, National Institute of Science and

Technology Policy (NISTEP) 984-1001

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 Trajtenberg, Manuel (1998), "International knowledge flows: Evidence from patent citations", NBER Working Paper No. 6507

Jaffe, Adam, and Manuel Trajtenberg (2003), *Patents, Citations and Innovations*, Cambridge: MIT Press

Jaffe, Adam, Manuel Trajtenberg and Fogarty (2000) *American Economic Review*

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

Keller, Wolfgang (1996), "Are international R&D spillovers trade-related? Analyzing spillovers among randomly-matched trade partners", SSRI Working Paper No. 9607, University of Wisconsin, Madison

Keller, Wolfgang (2002), "Geographic localization of international technology diffusion", *American Economic Review*, 92, pp 120-142

Keller, Wolfgang, and S. Yeaple (2003), "Multinational enterprises, international trade, and productivity growth: Firm level evidence from the United States", NBER Working Paper # 9504, February.

Keller, Wolfgang (2004), "International Technology Diffusion", forthcoming, *Journal of Economic Literature.*

Lichtenberg, Frank and Bruno van Pottelsberghe de la Potterie (2001) "Does foreign direct investment transfer technology across borders?," *The Review of Economics and Statistics*, 83(3), 490-7.

Levinsohn, Steve and Petrin, Amil (2003) "Estimating Production Functions using inputs to control for unobservables" *Review of Economic Studies*, 70, 317-341

Manski, Charles (1991), "Identification of endogenous social effects: The reflection problem", *Review of Economic Studies* 60 (3), pp 531-42

Marshak, J. and Andrews,W. C. (1944), "Random Simultaneous Equations and the theory of production" *Econometrica*, 62(1), 157-180

Olley, Steve and Pakes, Ariel (1996) "The dynamics of Productivity in the Telecommunications equipment industry" *Econometrica* 64(6), 1263-1297

OECD (2003), "Structural Analysis Data Base", Paris

OECD (2002), "Analytical Business Enterprise Research and Development", Paris

Serapio, M. G. Jr. and Dalton, D. H. (1999), "Globalization of industrial R&D: an examination of foreign direct investments in R&D in the United States", Research Policy 28, pp 303-316

Singh, J. (2003), "Knowledge diffusion and the role of multinational subsidiaries", Harvard Business School and Department of Economics, mimeo

Vernon, R. (1966) 'International investment and international trade in the product cycle' *Quarterly Journal of Economics*, LXXX, 190-207.

von Zedtwitz, M. and Gassmann, O. (2002), "Market versus technology drive in R&D internationalisation: four different patterns
of managing research and development", *Research Policy* 31, pp 569-588

A. Appendix: Data

In order to implement our empirical strategy we need to measure three types of information: the location of firms' innovative activity, firms' productivity performance, and the domestic and foreign spillover pools available to firms. We use data from the US Patent Office, firm accounts data, and OECD data on industry level R&D expenditure.

A.1. Innovative activity

The first dataset is the NBER patent citations data file which contains computerised records of over two million patents granted in the USA between 1901 and 1999. This is the largest electronic patent dataset in the world. We restrict our attention to patents applied for after 1975 as on citations are only available for patents applied for after this date. This is combined with firm accounting data from the Datastream on-line service which contains information on output, employment, investment and R&D.[35]

[35] More details of the matching between the datasets can be found in Bloom and Van Reenen (2000).

A.1.1. Inventor location

Patents identify the address (including country) of the inventor(s) listed on the patent application. Table 1 (in the main text) shows the primary inventor's country for the 63,733 patents matched to the 266 UK firms (i.e. those listed on the London Stock Exchange). For comparison, the final column lists the share of the primary inventor's country for the entire patent database of all patents registered in the USA between 1975 and 1998 (more than 2 million patents). As expected the share of UK inventors is much higher for the patents owned by the 266 UK firms (31.0% in column (2)) than for the whole sample of patents (3.0% in column (3)). Nevertheless, the US has the highest share of inventors even for the patents owned by the 266 UK firms (45.1%). The high share of patents owned by the 266 UK firms, but invented in the USA, is probably partly due to home-country bias from using a US dataset, but also reflects the county's strong innovative performance and the location of many UK firms in the USA. An overall bias towards US based patents should not bias our results as long as it is not different across firms in a way that is related to other firm characteristics.

A.1.2. Patent Citations

We also use data on patent citations to refine our measures of the location of firms' innovative activity. We assume that a patent owned by a UK firm, but invented by an inventor located in the USA, is more likely to be associated with technology sourcing behaviour if it cites other patents whose inventors were located in the USA. In particular, if a patent owned by a UK firm but invented by an inventor located in the USA does not cite any other patents whose inventors were located in the USA, this suggests that the patent is unlikely to be associated with technology sourcing. Such a patent is more likely to be associated with other motivations for locating R&D abroad, such as adapting existing technologies to the local market.

The 63,733 patents matched to our 266 UK firms make 472,998 citations to other patents, an average of 7.4 citations made by each patent. Of these 472,998 citations, 405,788 have information on the country location of the cited inventor. Because we are interested in whether firms are benefitting from external knowledge that has not been generated within the same firm we exclude self-citations, where a patent cites another patent that is owned by the same firm. 8.7% of all citations are made to patents owned by the same patenting subsidiary (or "assignee"), while a further 1.1% of all citations are made to a different assignee that is nevertheless part of the same parent firm.

Table A1 shows a cross-tab of the location of the citing and cited inventor for the 359,265 non-self citations in our sample. It is important to remember that all of these citations were made by patents that are owned by UK firms, even if the inventor was located in the US. Only 6.9% of citations made by UK inventors are made to another UK inventor, while 60.7% are made to a US inventor. In contrast, 71.8% of citations made by US inventors are made to other US inventors, while only 3.2% are made to UK inventors. This probably illustrates both the fact that the data is from the US patent office, but also the dominant global position of the USA in innovation. This provides preliminary evidence that most patents owned by UK firms but invented by an inventor located in the US are building on other knowledge created in the USA. When we look at self-citations to a patent that is owned by the same parent firm (not shown) the percentages in the diagonals (for example a UK inventor citing another UK inventor) are much higher. We also see that, even within firms, the transfer of knowledge from the UK to the USA appears to be small compared to the transfer of knowledge within the USA.

A.1.3. Patent Application dates

Geographic proximity is generally thought to be more important for the flow of knowledge that is "tacit", in the sense that it is not easily codified or written down in manuals. The flow of tacit knowledge is more likely to be mediated through face-to-face meetings and personal interactions between scientists and/or engineers. It also seems likely that knowledge that has been created recently is more "tacit" than knowledge that was created longer ago. Thus, firms that locate innovative activity in the US in order to gain access to pools of tacit knowledge are unlikely to be attempting to access knowledge that was created twenty or even ten years ago. For this reason we also use information on the application dates of each citing and cited patent in order to refine our measures of the location of firms' innovative activity. In particular we look at citations made to patents that were applied for within the last three years. Table A2 shows the same cross-tab of the country of the citing and cited inventor for all non self-citations of this type. The proportions are similar to those in Table A1, although UK inventors are slightly more likely to cite other UK inventors than before, while US inventors are less likely than before to cite other US inventors.

A.2. Firm Accounts data

We sought to construct similar types of data in both US and UK, although some differences were inevitable. The samples were both independently matched to USPTO data. The samples are based on publicly listed firms, whose primary sales are in manufacturing and who report some R&D between 1990 and 2000. Data relates to the consolidated worldwide accounts. Observations with missing data, firms with less than five consecutive observations over 1990 - 2000, and firms for which there were jumps greater than 150% in any of the key variables (capital, labour, sales) were deleted. R&D capital stocks were constructed by the perpetual inventory method using a depreciation rate of 15%.

A.2.1. UK firms

The data on value-added, labour (DS Item 182), capital and R&D expenditure (DS Item 119) comes from the Datastream On-Line service and is a sample of firms listed on the London Stock Exchange. Although these are "UK firms" in the sense that they are listed on the London Stock Exchange, a key feature of the data is that it relates to the firm's global activities.

Value added is the sum of total employment costs, operating profits, depreciation and interest payments.

The initial sample is all firms existing in 1985 with names starting with the letters A-L, plus any of the top 100 UK R&D performers not already included. The sample includes 415 firms, 266 of whom had taken out at least one patent between 1975 and 1998. All these firms' subsidiaries were located using *Who Owns Whom* by Dun and Bradstreet in 1985. All the subsidiaries were then matched by name to the USPTO (see Nick Bloom and John Van Reenen, 2002, for details)

The data does not include intermediate inputs, so value added was constructed as the sum of total employment costs, operating profit, depreciation and total interest charges. Most UK firms did not report R&D expenditure before 1989 and so the analysis is restricted to the years 1990-2000.[36] An R&D capital stock was constructed using a perpetual inventory method and an assumed 15% rate of obsolescence.

After cleaning our data we have a sample with 1794 observations on 188 firms, 141 of which are matched to at least one patent. Table 2 in the main text reports summary statistics. On average, firms in our sample have applied for 240 patents. 46.2% of these patents are taken out by inventors located in the US, 41.7% are taken out by inventors located in the US and also cite at least one other US patent, and 16.2% are taken out by inventors located in the US, cites at least one other US patent and that has been applied for within the last three years.

[36] Even after 1989 when a firm reports zero R&D it is not clear that this corresponds to a true zero, although it is unlikely to perform a large amount of R&D. In the results presented in this paper, a dummy variable was used to denote reported zero R&D expenditure, but the results are not sensitive to the exact treatment of reported zeros.

A.2.2. US firms

US Data was taken from the match between Compustat and the USPTO conducted by Bronwyn Hall et al (2000). We tried to make the sample and variable construction as close as possible to the UK sample. We matched in industry level data by primary SIC code as above (1987 Revision). The book value of capital is the net stock of property, plant and equipment (CS Item A8 - PPENT). R&D is CS item A46 - XRD. Unfortunately staff costs are only available for 5% of firms in the Compustat data so constructing a value added measure is extremely difficult. Consequently we follow the tradition in the US literature and use real sales as our output measure (CS Item A12- SALE).

The inventors of US firms are much more localised in the United States than in UK firms (see Table 1). 95% of all inventors were in the US and only about 1% of inventors were located in the UK. This reflects the innovative strength of the US and the fact we are using USPTO data, so there is some inevitable home bias for the US. The industries where there is greater US innovative presence in the UK are (unsurprisingly) those were the UK has some traditional strengths - medical equipment, pharmaceuticals, and petroleum refining.

Table A4 describes the data on US firms.

A.3. Industry level data - R&D Spillover pool

The domestic and foreign spillover pools were constructed using the OECD's Analytical Business Expenditure on R&D dataset (ANBERD, 2002). This contains information on R&D spending at the 2-digit manufacturing industry (ISIC Revision 3) for all OECD countries. A stock measure was constructed using a perpetual inventory method and an assumed 15% rate of obsolescence,[37] with a starting year of 1987. Although there are various problems with using industry-level measures, as discussed above, this data has the crucial advantage for our purposes that it contains R&D expenditures by geographical location of the R&D activity. This would be extremely hard to recreate using data on firms' R&D as they do not decompose R&D into a foreign and domestic element. Our measure also has the advantage of including all R&D carried out in each industry in each country, and not just the R&D of the other sampled firms. We also use data on 2-digit industry level value-added taken from the OECD's Structural Analysis database (STAN, 2003). Value added price deflators at the two digit level are also used from this source.

[37] We experimented with other depreciation rates but the results were not significantly changed.

B. Appendix: Econometric modelling strategy

We present baseline OLS production function of equation (2.6) but there are several well known problems with obtaining consistent estimates of the parameters. These relate to the endogeneity of the inputs, unobserved heterogeneity and survivor bias/selectivity. We compare results from two alternative approaches to deal with these problems, a GMM method (Richard Blundell and Stephen Bond, 2000) and the popular "OP" method (Stephen Olley and Ariel Pakes, 1996) adapted for the presence of endogenous R&D (Thomas Buettner, 2004).

B.1. System GMM

Consider a simplified form of the production function

$$y_{it} = \alpha x_{it} + a_{it} \tag{B.1}$$

where x_{it} is an endogenous input and the residual productivity term takes the form

$$a_{it} = t_t + \eta_i + u_{it}. \tag{B.2}$$

Year dummies (t_t) control for common macro effects and the firm fixed effect (η_i) and stochastic productivity shock (u_{it}) may be correlated with the regressors. Assuming no serial correlation in the u_{it} process yields the following moment conditions

$$E[x_{i,t-s}\Delta u_{it}] = 0 \tag{B.3}$$

for $s \geq 2$[38]. This allows the use of suitably lagged levels of the variables to be used as instruments after the equation has been first differenced. We test for the problem of serial correlation through LM tests at the base of all the GMM columns. If there is higher order (but finite) serial correlation in the u_{it} process longer lags can still be used as instruments.

Unfortunately, the first differenced GMM estimator has been found to have poor finite sample properties when the endogenous variables are highly persistent because the lagged instruments are weakly correlated with the first differences of the endogenous variables. If we are prepared to make an initial conditions

[38] If there is serial correlation in the error term this can be dealt with by using longer lags as instruments. For example, if $u_{it} \sim MA(1)$ lags dated t$-$3 and earlier will be valid instruments.

assumption $\mathrm{E}[\Delta y_{i2}\eta_i] = 0$ and $\mathrm{E}[\Delta x_{it}\eta_i] = 0$ then additional moment conditions become available[39]. The additional moment conditions take the form:

$$E[\Delta x_{i,t-s}(\eta_i + u_{it})] = 0 \tag{B.4}$$

for $s = 1$ when $u_{it} \sim MA(0)$. This means that the lagged difference of x can be used as instruments in the levels equations. We test the validity of the additional moment conditions using a Sargan difference test. The levels equations and differenced equations are stacked in a system each with its appropriate instruments.

We assume that all firm-level variables are endogenous, whereas all industry-level variables are treated as exogenous. We examine specifications where the industry-level R&D stocks are treated as endogenous and the results are not significantly affected. The results are also robust to lagging the industry-level variables by one period, in which case they can be treated as pre-determined. We instrument firm-level variables in the differenced equation with their levels lagged from two to five times inclusive, and in the levels equation by their first-differences lagged once, as well as by all time and industry dummies and all exogenous variables.

The standard errors we present allows for arbitrary heteroskedasticity and arbitrary serial correlation. They are the "One-Step robust" results from the DPD package written in GAUSS[40] (i.e. we do not iterate on the GMM weight matrix because of the Monte Carlo evidence of underestimation of the second step standard errors). We include full sets of time dummies and industry dummies in all regressions.

B.2. Olley Pakes with endogenous R&D

Olley and Pakes (1996) essentially assume that the production function can be written

$$y_{it} = \alpha_0 + \alpha_l l_{it} + \alpha_k k_{it} + \omega_{it} + \upsilon_{it} \tag{B.5}$$

where ω_{it} is the productivity state and υ_{it} is a serially uncorrelated additional productivity shock or measurement error (which can be serially correlated). This

[39] Stationarity of y_{it} and x_{it} is sufficient (but not necessary) for these conditions to hold. What is essential is that teh first moments of the endogenous variables are time invariant conditional on the time dummies. The higher order moments are unrestricted.

[40] Available from: http://www.ifs.org.uk/econometindex.shtml

is equation (2.4) with $\beta = \gamma_{i1} = \gamma_{i2} = 0$ and $a_{it} = \omega_{it} + \upsilon_{it}$.Capital is quasi-fixed and labour completely variable. The bones of the Olley Pakes model is as follows. At the beginning of the period t, firm i observes its productivity state ω and capital k. The key difference between ω and υ is that ω is a state variable and affects invetsment decisions whereas υ does not. The firm decides whether to stay in business based on its expectations of net present value value compared to a critical cut off. Denote $\chi_t = 1$ if the firm chooses to stay in business and $\chi_t = 0$ if the firm chooses to exit. If the firm decides to continue operations it sets labour and chooses the level of investment in physical capital. Physical capital evolves in a deterministic process based on investment according to the standard perpetual inventory formula. The additional shock υ is then realized after these choices are made. The key insight of OP algorithm is to use the monotonicity of the investment policy function in unobserved productivity (conditional on current capital). This can be used to get consistent estimates of the parameter on variable inputs at stage 1 and then use these (at stage 2) to obtain the capital coefficient.

We take two approached to dealing with firm R&D. First, we consider estimates of the standard OP algorithm and include R&D as an exogenous variable[41]. Secondly, we follow Thomas Buettner's (2004) extension of the OP structural model to include endogenous R&D chosen at the same time as fixed investment. Unlike fixed investment, however, R&D is stochastic. The productivity state ω still evolves stochastically over time according to a controlled Markov process, but the distribution of next period's productivity is increasing (in a first order stochastic dominance sense) not only in the current productivity state *but also in the amount of R&D expenditure.* We can think of this as the firm "buying" an improved probability distribution of ω_{t+1} through spending more on R&D this period. We assume that the distribution of ω_{t+1} is governed by a parameter $\psi_{t,}$ a single index. The distribution of next period's productivity ω_{t+1} is a member of the family of distributions.[42]

$$\digamma_{\psi_{t+1}} = \{F(\omega_{t+1}|\psi_{t+1}), \psi_{t+1} \in \Psi\}$$

An important contribution of Buettner is to show that (in the context of this extended structural model) the invertibility of the investment policy function still holds. Consequently stage 1 of the OP algorithm does not need to be changed.

[41] Analagously to plant age in the original Olley Pakes (1996) application.

[42] This is an important restriction as it implies that R&D and ω_{it} affects ω_{it+1} only through $\psi_{it+1.}$ Thus productivity shocks and R&D are not allowed to have a qualitively different impact on the distribution of future productivity.

B.2.1. Stage One: Estimation of the coefficient of the variable input.

The estimation strategy is to control for the unobserved productivity shock non-parametrically by exploiting the monotonicity of the investment policy function. Unobserved productivity can be written as[43]

$$\omega_{it} = \widetilde{\omega}(i_{it,}k_{it})$$

Substituting this into the production function (B.5) gives

$$y_{it} = \alpha_0 + \alpha_l l_{it} + \phi(i_{it,}k_{it}) + \upsilon_{it} \tag{B.6}$$

where

$$\phi_t = \phi(i_{it,}k_{it}) \equiv \alpha_0 + \alpha_k k_{it} + \widetilde{\omega}(i_{it,}k_{it})$$

We do not know the functional form of ϕ_t so we use a series estimator to approximate it [44]. Estimation of equation (B.6) gives a consistent estimate of α_l and estimates of the unknown function ϕ_t.

B.2.2. Stage Two: Estimation of the coefficient of the quasi-fixed input.

Rearranging (B.6) after we have an estimate of the coefficient on the variable input (α_l) gives

$$y^*_{it} = y_{it} - \alpha_l l_{it} = \alpha_0 + \alpha_k k_{it} + \upsilon_{it}$$

The expectation of y^*_{it} conditional on information at $t-1$ and survival until t is then

$$E[y^*_{it}|I_{t-1}, \chi_{it} = 1] = \alpha_0 + \alpha_k k_{it} + E[\omega_{it}|\psi_{it,}\chi_{it} = 1]$$

where I_{t-1} is the information set in $t-1$,the distribution of productivity states is $\psi_{it,}$ (which is influenced by the firm's R&D choice). Under the Markov assumption for productivity, we can re-write productivity conditional on survival as:

$$\omega_{it} = E[\omega_{it}|\psi_{it,}\chi_{it} = 1] + \xi_{it}$$

The second stage estimation becomes

[43] Or equivalently $\widetilde{\omega}(k_{it+1,}k_{it})$ since capital is formed deterministically: $k_{it+1} = (1-\delta)k_{it}$.

[44] Olley and Pakes find that the fully non-parametric estimator of ϕ_t gives similar results to the series estimator. We found that fourth or sixth order series expansions (instead of our preferred fifth order) made little difference to the results.

$$y_{it}^{*}=\alpha_0+\alpha_k k_{it}+E[\omega_{it}|\psi_{it,}\chi_{it}=1]+\xi_{it}+\upsilon_{it}$$

where the productivity innovation ξ_{it} is uncorrelated with k_{it}.To control for selectivity we will take a similar approach to stage 1 and control for the expectation non-parametrically.

In the absence of selection and R&D the second stage becomes simply

$$y_{it}^{*}=\alpha_0+\alpha_k k_{it}+g(\omega_{it-1})+\xi_{it}+\eta_{it} \tag{B.7}$$

Since $\omega_{it-1}=\phi_{t-1}-\alpha_k k_{it-1}-\alpha_0$, equation (B.7) can be estimated by non-linear least squares where the unknown function $g(\omega_{it-1})$ can be approximated by a nonparametric function in $\phi_{t-1}-\alpha_k k_{it-1}$.The key difference between Buettner's model and the original OP model is that $\psi_{it,}$ depends on both ω_{it-1} and k_{it-1} in the model with endogenous R&D whereas it only depends on ω_{it-1} in the original OP set-up. This means that there is a difference in the method in which we estimate stage 2. We use the fact that the R&D function can be written $r(\psi_{it,}\omega_{it-1})$ and invert this to obtain

$$\psi_{it}=r^{-1}(r_{it-1,}\omega_{it-1}) \tag{B.8}$$

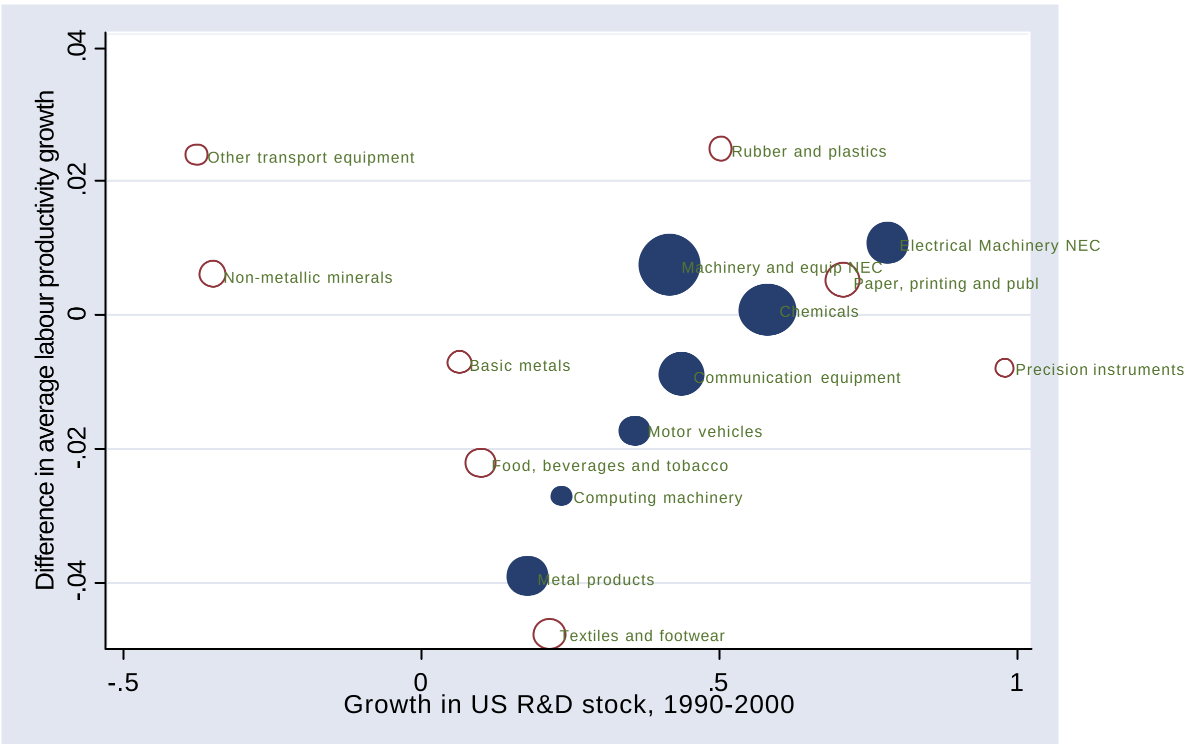
where $r_{it-1,}$ denotes the observed R&D spend at t-1.Using equation (B.8) to control for the distribution in period t, the second stage estimation equation becomes

$$\begin{aligned} y_{it}^{*} &= \alpha_k k_{it}+g(r^{-1}(r_{it-1,}\omega_{it-1}))+\xi_{it}+\upsilon_{it} \\ &= \alpha_k k_{it}+\widetilde{g}(r_{it-1,}\phi_{t-1}-\alpha_k k_{it-1})+\xi_{it}+\upsilon_{it} \end{aligned} \tag{B.9}$$

Equation (B.8) can be used to obtain estimates for α_k replacing $g(r^{-1}(.,.))$ with a nonparametric function $\widetilde{g}(.,.)$ in r_{it-1} and $\phi_{t-1}-\alpha_k k_{it-1}$.

Armed with these estimates for the parameters of the production function we can then construct the productivity term ω_{it}. These estimates of the productivity term are then related to the spillover terms in a final stage regression.

Figure 1: US R&D growth and "productivity growth premium" for UK firms with a high proportion of US inventors



Notes:

Vertical axis is the "productivity premium" for UK firms with strong inventor presence in the US between 1990 and 2000 (i.e. the differential in annual average labour productivity growth for our UK firms with above median US inventor presence versus those with below median US inventor presence). The horizontal axis is total growth in US R&D stock. Shaded industries are those with largest US-UK TFP gap in 1990 (i.e. where UK firms had the "most to learn"). Industry points are weighted by number of firms in our sample. Although there is a positive relationship across all industries, it is strongest in the "high gap" sector.

Table 1: Country of inventor

Country of Inventor	(1) Number of Patents matched to our UK firms	(2) % Share of patents matched to our UK firms	(3) % Share of patents matched to our US firms	(4) % Share of all USPTO patents
UK	19,745	31.0	1.1	3.0
USA	28,731	45.1	92.3	55.7
Japan	4,411	6.9	1.5	18.8
Germany	2,481	3.9	1.3	7.9
France	1,457	2.3	0.9	3.0
Other	6,908	10.8	2.9	11.6
Total	63,733	100	100	100

Notes: First two columns give inventor location from the matching of UK firms to USPTO. Column (3) from matching of US firms to USPTO. Final column refers to all patents registered at the US Patent Office between 1975 and 1998

Table 2: Descriptive Statistics for UK firms

	Mean	Median	Standard Deviation
Employees	10,711	1,750	27,564
Value added (£m)	372	48	914
Capital stock per worker (£)	38,700	30,000	31,900
Value added per employee (£)	31,404	50,201	12,438
R&D expenditure/value added	0.029	0.010	0.044
R&D stock/value added	0.158	0.046	0.272

Notes: 188 firms, 1990-2000; all monetary amounts are in 1995 currency, deflated using OECD 2 digit industry price deflator; value added is constructed as the sum of total employment costs, operating profit, depreciation and interest payments; capital stock and R&D stock are constructed using a perpetual inventory method as described in the text

Table 3: R&D-Augmented Production Functions

	(1)	(2)	(3)	(4)	(5)	(6)
Estimation Method	**OLS**	**OLS**	**GMM**	**GMM**	**GMM**	**Olley-Pakes**
Dependent variable	**ln (Y)** $_{it}$	**ln (Y/K)** $_{it}$	**ln (Y/K)** $_{it}$	**ln (Y/K)** $_{it}$	**ln (Y/K)** $_{it}$	**ln (Y)** $_{it}$
Company listed in:	**UK**	**UK**	**UK**	**UK**	**UK**	**UK**
Location weight: W_i	Location	Location	Location	Location & Citation	Location & Citation within 3 years	Location & Citation within 3 years
ln (L/K) $_{it}$ **labour-capital**	-	0.657 (0.046)	0.648 (0.065)	0.647 (0.065)	0.642 (0.067)	-
ln (L) $_{it}$ **labour**	0.620 (0.057)	-	-	-	-	0.575 (0.041)
Ln(K) $_{it}$ **capital**	0.343 (0.042)	-	-	-	-	0.432 (0.045)
ln (R&D) $_{it,}$ **firm R&D stock**	0.029 (0.008)	0.012 (0.007)	0.026 (0.011)	0.025 (0.010)	0.022 (0.010)	0.008 (0.004)
W_i^{US} *** ln (US R&D)** $_{jt}$	-	0.076 (0.024)	0.066 (0.035)	0.084 (0.031)	0.173 (0.054)	0.115 (0.045)
W_i^{UK} *** ln (UK R&D)** $_{jt}$	-	0.035 (0.022)	0.026 (0.028)	0.092 (0.095)	0.400 (0.291)	0.147 (0.338)
ln (US R&D) $_{jt}$ **US industry R&D stock**	-	0.050 (0.118)	0.065 (0.067)	0.059 (0.065)	0.063 (0.066)	-0.050 (0.030)
ln (UK R&D) $_{jt}$ **UK industry R&D stock**	-	0.273 (0.165)	0.221 (0.101)	0.219 (0.101)	0.206 (0.096)	0.120 (0.053)

W_i^{US} **% inventors in US**	-	-0.696 (0.240)	-0.602 (0.336)	-0.765 (0.313)	-1.658 (0.543)	-1.097 (0.439)
W_i^{UK} **% inventors in UK**	-	-0.296 (0.156)	-0.254 (0.193)	-0.760 (0.683)	-3.270 (2.533)	-1.280 (2.506)
Firms	188	188	188	188	188	188
Observations	1794	1794	1794	1794	1794	1794
1st order serial correlation test *(p-value)*	-	-	-1.212 *(0.226)*	-1.212 *(0.226)*	-1.212 *(0.225)*	-
2nd order serial correlation *(p-value)*	-	-	-1.788 *(0.074)*	-1.769 *(0.077)*	-1.719 *(0.086)*	-
Sargan Difference Test *(p-value)*			17.52 *(0.562)*	17.90 *(0.534)*	18.81 *(0.456)*	-
Sargan Test of Over-identifying restrictions *(p-value)*	-	-	86.39 *(0.217)*	86.18 *(0.222)*	86.52 *(0.214)*	-

Notes:

W_i^{US} and W_i^{UK} are the (pre-1990) proportion of a firm's inventors located in the US and UK respectively. Standard errors in brackets under coefficients are robust to heteroskedacity and autocorrelation of unknown form and are clustered by industry. The dependent variable in columns (2) through (5) is the log of value added divided by capital stock, in column (6) it is the log of value added and in column (7) it is the log of real sales. The time period is 1990-2000. Columns (1), (2) and (7) are estimated by OLS. Columns (3) to (5) are estimated by System-GMM (one-step robust standard errors). In Systems GMM (see Blundell and Bond, 2000) the firm-level variables are assumed endogenous and industry level variables are assumed strictly exogenous; endogenous variables are instrumented by levels lagged from two to five times in the differences equation and differences lagged once in the levels equation, as well as by all exogenous variables and year and industry dummies. Column (6) is estimated by the OP method (Olley-Pakes, 1996). In OP we use a fifth order series expansion in the first and second stage (the second stage also includes a selection correction term). After obtaining the firm specific (total factor) productivity term (? $_{it}$) from stage one we regress this against the indicated variables (including full sets of industry and time dummies). In OP the standard errors are bootstrapped (100 replications) and allow for clustering by firm. For diagnostic tests p-values are in brackets and italics. Columns (1) through (6) are UK firms and column (7) is the sample of US firms. All equations include a full set of industry dummies and time dummies. Column (7) also includes US industry value added (which was insignificant in the other columns).

Table 4: R&D Augmented Production function results – Further Investigations

	(1)	(2)	(3)	(4)
Estimation method	**GMM**	**GMM**	**GMM**	**GMM**
Dependent variable	**ln (Y)** $_{it}$	**Log(Y/K)** $_{it}$	**Log(Y/K)** $_{it}$	**Log(Y/K)** $_{it}$
Company listed in	**USA**	**UK**	**UK**	**UK**
Sample	USA	**“Domestic”**	**High TFP Gap with USA**	**Low Gap with the USA**
Location weight:	Location	Location & Citation within 3 years	Location & Citation within 3 years	Location & Citation within 3 years
ln (L/K) $_{it}$	-	0.610	0.757	0.518
		(0.072	(0.076)	(0.087)
ln (L) $_{it}$	0.706			
	(0.078)			
ln (K) $_{it}$	0.220			
	(0.052)			
ln (R&D) $_{it}$	0.049	0.029	0.029	0.005
	(0.035)	(0.014)	(0.013)	(0.014)
W_i^{US} *** ln (US R&D)** $_{jt}$	0.002	0.212	0.277	0.123
	(0.072)	(0.063	(0.130)	(0.093)
W_i^{UK} *** ln (UK R&D)** $_{jt}$	0.151	-0.672	0.434	-0.826
	(0.131)	(0.408	(0.267)	(1.072)
ln (US R&D) $_{jt}$	0.247	0.116	0.353	0.035
	(0.078)	(0.096	(0.171)	(0.070)
ln (UK R&D) $_{jt}$	-0.063	0.211	0.404	-0.041
	(0.046)	(0.115	(0.152)	(0.121)

W_i^{US}	-1.244 (0.978)	-2.028 (0.637)	-2.849 (1.445)	-1.182 (0.844)
W_i^{UK}	-0.097 (0.781)	4.199 (2.757)	-3.540 (2.338)	4.861 (7.040)
Firms	570	135	99	89
Observations	5446	1267	938	856
1st order serial correlation (p-value)	*-4.877* *(0.000)*	-1.198 (0.231)	-1.101 (0.271)	-2.702 (0.007)
2nd order serial correlation	-1.739 *(0.082)*	-1.814 (0.070)	-0.243 (0.808)	-1.468 (0.142)
Sargan difference test		13.99 (0.693)		
Sargan	67.96 *(0.081)*	83.63 (0.283)	55.22 (0.801)	66.93 (0.197)

Notes:

"Domestic" indicates the sub-sample of UK firms who are estimated to have no overseas production facilities. "High TFP Gap" indicates those industries where the TFP gap with the USA was above the median (see Figure 1). Column (1) contains US firms and columns (2) through (4) contain UK firms. W_i^{US} and W_i^{UK} are the (pre-1990) proportion of a firm's inventors located in the US and UK respectively. Standard errors in brackets under coefficients are robust to heteroskedacity and autocorrelation of unknown form. The dependent variable in columns (2) through (4) is the log of value added divided by capital stock and in column (1) it is the log of real sales. The time period is 1990-2000. All columns are estimated by System-GMM (one-step robust standard errors). The firm-level variables are assumed endogenous and industry level variables are assumed exogenous. Endogenous variables are instrumented by levels lagged from two to five times in the differences equation and differences lagged once in the levels equation, as well as by all exogenous variables and year and industry dummies. For diagnostic tests p-values are in brackets and italics. All equations include a full set of industry dummies and time dummies. Column (1) also includes US industry value added (which was insignificant in the other columns).

Table 5: Citations results

	(1)	(2)	(3)	(4)
Estimation method	**Negative Binomial**	**Negative Binomial**	**Probit**	**Probit**
Dependent variable	$CITES_{pit}^{US}$	$CITES_{pit}^{US}$	$1(CITES_{pit}^{US} > 0)$	$1(CITES_{pit}^{US} > 0)$
W_i^{US}	0.631	0.104	0.126	0.033
	(0.267)	(0.198)	(0.043)	(0.044)
W_i^{UK}	0.197	0.054	0.056	0.025
	(0.205)	(0.199)	(0.042)	(0.047)
US_{pit}	-	0.684	-	0.124
		(0.158)		(0.020)
UK_{pit}	-	0.037	-	0.010
		(0.107)		(0.023)
$TOTALCITES_{pit}$	0.013	0.012	0.009	0.008
	(0.001)	(0.001)	(0.002)	(0.002)
Dispersion (delta)	1.050	0.999	-	-
	(0.069)	(0.067)		
Observations	14,161	14,161	14,161	14,161
Mean of dep. var.	0.695	0.695	0.356	0.356
Log Pseudo-L	-15,116.06	-14,996.25	-8,465.02	-8,400.51

Notes: The dependent variable in columns (1) and (2) is the number of citations per patent to a US inventor (not owned by the same firm and applied for within the last three years); the dependent variable in columns (3) and (4) is a dummy equal to one if the patent cited at least one patent with a US inventor (not owned by the same firm and applied for within the last three years). Reported coefficients in columns (1) and (2) are equal to the incidence-rate ratio minus one, and in columns (3) and (4) are marginal effects. W_i^{US} and W_i^{UK} are the (pre-1990) proportion of a firm's inventors located in the US and UK respectively. US_{pit} and UK_{pit} denote whether the patent's inventor is located in the US or UK respectively. *TOTALCITES* is the total number of cites made by the patent. Robust standard errors in brackets are adjusted for clustering by firm; all specifications include 8 year dummies, 14 industry dummies and 36 technology class dummies, as well as all firm and industry level variables from the production function in Table 3. The sample consists of all patents applied for by our UK firms between 1990 and 1998.

Appendix Tables

Table A1: Location of citing and cited inventors: non self-citations

Cited country:	UK	USA	Other	Total
Citing country:				
UK	5,842	51,370	27,427	84,639
	(6.9%)	(60.7%)	(32.4%)	(100%)
USA	6,445	145,180	50,529	202,154
	(3.2%)	(71.8%)	(25.0%)	(100%)
Other	2,387	39,684	30,401	72,472
	(3.3%)	(54.8%)	(42.0%)	(100%)
Total	14,674	236,234	108,357	359,265
	(4.1%)	(65.8%)	(30.2%)	(100%)

Notes: Citations made by the firms in our sample, excluding self-citations.

Table A2: Location of citing and cited inventors: non self-citations, cite at least one patent with US inventor that has been applied for within past three years

Cited country:	**UK**	**USA**	**Other**	**Total**
Citing country:				
UK	717	3,697	5,140	9,554
	(7.5%)	(38.7%)	(53.8%)	(100%)
USA	224	4,397	3,847	8,468
	(2.7%)	(51.9%)	(45.4%)	(100%)
Other	512	5,167	12,610	18,289
	(2.8%)	(28.3%)	(69.0%)	(100%)
Total	1,453	13,261	21,597	36,311
	(4.0%)	(36.5%)	(59.5%)	(100%)

Notes: Citations made by the firms in our sample.

Table A3: Summary statistics for UK patenting firms

	Mean	Median	Standard Deviation	Min	Max
Total patent applications	240	40.5	657	1	5820
UK Location Weight	0.354	0.274	0.363	0	1
UK Location + Citation Weight	0.082	0.017	0.145	0	1
UK Location + Citation Within 3 Years	0.019	0.000	0.054	0	0.5
USA Location Weight	0.462	0.425	0.379	0	1
USA Location + Citation Weight	0.417	0.368	0.349	0	1
USA Location + Citation Within 3 Years	0.162	0.134	0.184	0	1

Notes: 141 firms matched to at least one patent; location weights are constructed as described in the text

Table A4 Descriptive Statistics for US firms

	Mean	Median	Standard Deviation
Employees	13,760	3,528	38,640
Real Sales ($1000)	3,196	586.4	10,742
Capital per employee ($)	59,407	34,607	81,630
Real sales per employee ($1000s)	193.736	162.843	128.641
R&D expenditure/value added	0.059	0.029	.198
R&D stock/value added	0.237	0.113	0.567

Notes: All in 1995 prices, 570 firms, 6016 observations, 1990-2000