Spillovers in Space:
Does Geography Matter?

Sergey Lychagin and Joris Pinkse*
Center for Auctions, Procurements and Competition Policy
Department of Economics
The Pennsylvania State University

Margaret E. Slade†
Department of Economics
The University of British Columbia

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

July 2010

ABSTRACT

We simultaneously assess the contributions to productivity of three sources of research and development spillovers: geographic, technology and product–market proximity. To do this, we construct a new measure of geographic proximity that is based on the distribution of a firm's inventor locations rather than its headquarters, and we report both parametric and semiparametric estimates of our geographic–distance functions. We find that: i) Geographic space matters even after conditioning on horizontal and technological spillovers; ii) Technological proximity matters; iii) Product–market proximity is less important; iv) Locations of researchers are more important than headquarters but both have explanatory power; and v) Geographic markets are very local.

Keywords: R&D spillovers, geographic proximity, technological proximity, semiparametric estimation, spatial econometrics

JEL Classifications: O33, L60, C23

*Joris Pinkse's research is supported by NSF grant SES-0922127. Sergey Lychagin and Joris Pinkse thank the Human Capital Foundation for their support of CAPCP.

†Corresponding author: m.slade@mac.com. Margaret Slade thanks the Social Sciences and Humanities Research Council of Canada for support.

‡John Van Reenen would like to thank the Economic and Social Research Council for financial support through the Centre for Economic Performance.

1 Introduction

Firms engage in research and development (R&D) because they anticipate future benefits in the form of lower costs, new methods of production, and new uses for their products. To the extent that they can capture the fruits of their efforts, private and social returns to R&D should be equal.[1] However, when they cannot appropriate the returns, R&D becomes a public good and underinvestment is expected. A knowledge of the magnitude and direction of the problem (i.e., who benefits from the efforts of whom) is therefore a prerequisite for an appropriate design of public policy.

We address the much studied question of the relationship between R&D spillovers and productivity, where a spillover occurs when firm i benefits from the R&D activities of firm j. There is substantial evidence that spillovers exist. Furthermore, it is commonly agreed that spillovers are larger when firms are *closer*. However, it is not clear what closeness means or how it can be measured, and those questions are the focus of our research. Specifically, although we are most interested in geographic proximity, we look at variants of three measures of closeness: proximity in horizontal (product market), technological (R&D), and geographic space.[2]

Our research differs from the literature to date in three respects. First, we assess all three measures simultaneously, which removes the possibility that one measure (say geographic proximity) appears to be important simply because researchers fail to control adequately for the common influence of a second measure (say technological proximity). In other words, firms that are technologically similar might locate in the same region and benefit from each other's R&D activities. There has been little work that uses the same data and model to evaluate multiple measures of closeness, both individually and jointly.

Second, we refine the measures of closeness, particularly geographic proximity. To illustrate, most previous researchers have used the (single) location of a firm's headquarters to calculate geographic distance. It is not clear, however, that technological advances are likely to be communicated by CEOs. We hypothesize that inventors are more apt to communicate and we use the distribution of a firm's (multiple) inventor locations to calculate proximity. To our knowledge, no one has exploited firms' research locations to create a multidimensional measure of geographic closeness.

Third, whereas previous researchers have used either a zero/one measure of geographic distance (same or different region) or a simple one–parameter function of Euclidean distance between headquarters, we allow the effect of our multi–dimensional measure of geographic proximity to decay in a more flexible manner. Specifically, we estimate that

[1] For optimality, technology markets must also be competitive. In other words, strategic behavior can also distort decisions (see, e.g., Spence 1984).

[2] We do not assess vertical proximity because we do not have adequate data.

function semiparametrically.

Our research attempts to answer a number of questions. First and foremost, what is the nature of spillovers? In particular, who benefits from the R&D activities of whom? This question is logically prior to all others. Indeed, in the absence of an answer, it is impossible to formulate appropriate public policy towards R&D. Furthermore, it cannot be answered by considering one source of spillovers in isolation. Second, to what extent is knowledge private or public? In other words, can firms capture the lion's share of the rents that are associated with their R&D activities or are they freely available to rivals? This question must be answered if we are interested in assessing under and over investment and the need to correct the associated externalities. Finally, are knowledge flows confined to narrowly defined product markets, technology classes and/or geographic regions? Put differently, are spillovers local or global? This question must be answered if we are concerned with asymmetric patterns of regional or industrial growth and decline and whether growth paths will converge or diverge.

We use U.S. firm level accounting data (sales, employment, capital, R&D, etc.) from U.S. Compustat 1980-2000 matched into the U.S. Patent and Trademark Office (USPTO) data from the NBER data archive. The Compustat data were compiled by Bloom, Schankerman, and Van Reenen (2008). The use of patent data to create distributions of inventor locations, however, is new.

To analyze that data, we employ spatial econometric techniques that were developed by Pinkse, Slade, and Brett (2002) and Pinkse and Slade (2004). With those techniques, which allow us to handle multiple measures of closeness in a flexible but parsimonious manner, space is broadly defined to include a number of geographic and nongeographic dimensions.

Section 2 discusses the previous literature, section 3 the model, and section 4 measuring spillovers. The data is discussed in section 5, econometric methods in section 6, and results in section 7. Some concluding comments are offered in section 8.

2 Previous Literature

Griliches (1979) was perhaps the first to recognize the multi–dimensional nature of the spillover problem. He discussed alternative hypotheses about the nature of closeness, which can be:

Horizontal or learning from product–market rivals

Firms that produce similar products can often benefit from each other's R&D activities. For example, when a pharmaceutical firm introduces a new drug, in the absence of patent protection, rival companies can easily determine its makeup and offer close substitutes. Horizontal spillovers have been studied by e.g., Levin and Reiss (1988), Bernstein (1988), and Ornaghi (2006).

Technological or learning from technology–market rivals

Firms that perform similar types of research can also learn from each other. For example, the discovery of froth flotation, which facilitated the refining of sulfide ores, was an important advance in the mining industry that benefitted firms from many different product markets (e.g., copper, lead, and zinc metals, which are used in wiring, batteries, and galvanizing, respectively). Technology–market spillovers have been studied by e.g., Jaffe (1986).

Vertical or learning from suppliers or retailers

Firms that are related through a vertical chain might experience technological synergies. Although vertical or input/output spillovers have been studied, it is not clear if these are real or simply result from failure to control for quality or price. For example, if there is an advance that increases the quality or lowers the cost of producing auto bodies, auto manufacturers benefit. However, their technology has not changed; they can simply purchase better bodies at lower prices (a pecuniary externality). Vertical spillovers have been studied by e.g., Scherer (1982), Griliches and Lichtenberg (1984), and Goto and Suzuki (1989).

More recently, researchers have recognized that there is another type of closeness that Griliches (1979) did not mention:

Geographic or learning by meeting

Firms that are located in the same or proximate regions have opportunities to communicate. For example, employees of different firms that are located in close geographic proximity might meet on the golf course or belong to common civic organizations where they discuss and learn from each other's research activities. In other words, social networks can facilitate learning. In addition, geographic proximity can enable capitalization of complementarities among firms' research activities. Geographic spillovers are particularly important since they affect the potential for regional and national convergence. Indeed, if spillovers are highly localized, the prospects for convergence are poor. Geographic spillovers have been studied by e.g., Jaffe, Trajtenberg and Henderson (1993), Eaton and Kortum (1996), Keller (2002), and Orlando (2004).

Most early spillover studies used data on industrial sectors to examine intersectoral input/output flows or vertical spillovers based on sectoral input/output tables. More recently, however, researchers have used firm level data that allow them to examine such issues as intra versus inter product or geographic market spillovers. In general, researchers have found significant spillover effects. However, most concentrate on one proximity measure in isolation.[3] In addition, all use scalar measures of geographic proximity that depend on at most one parameter.

Our research, which is based on firm data in a multi–proximity–measure setting, takes a production–function approach to the problem. There are of course other approaches. For example, some researchers (e.g., Bernstein and Nadiri 1989) have used a dual framework based on a system of factor demands.[4] The advantage of the former is that it requires no behavioral assumptions and therefore avoids misspecification bias. The dual approach, in contrast, requires strong assumptions such as perfectly competitive input and output markets or, if imperfect competition is assumed, on the nature of the product market game or the bargaining process. Since the latter exploits more information, the estimates obtained are more efficient if that information is correct. Unfortunately, if that information is inaccurate, biases are introduced. Furthermore, since a system of interrelated equations is estimated, the bias in a single equation can contaminate all equations.[5]

There are a number of issues that we have chosen not to cover. For example, we consider private R&D but neglect public and academic activity (for work on the latter, see e.g., Adams 1990, Jaffe 1989, and Acs, Audretsch, and Feldman 1992). Furthermore, although we exploit patent data to create our measure of geographic proximity, we do not consider patent flows as direct measures of spillovers (as in, e.g., Jaffe, Tratjenberg, and Henderson 1993). Finally and most importantly, we do not consider vertical or input/output spillovers. We do this for two reasons. First, we do not have detailed data on interfirm purchases and sales. In addition, we believe that most such studies suffer from a measurement problem that is due to the fact that both price and quantity aggregates underestimate the value of new or higher quality goods. Productivity is therefore apt to be mismeasured.

Finally, as is typical of studies that use cross sectional or panel–data, we do not exploit information on individual industries but instead use a common model for all. We adopt a common framework because ultimately we want to test whether a simple model has wide applicability.

[3]Exceptions include Orlando (2004), who assesses technological and geographic proximity in a discrete framework (e.g., same product market/different geographic market), and Bloom, Schankerman and Van Reenen (2008), who assess product and technology–market proximity.

[4]Still others have used a consumer–surplus framework (see, e.g., the survey by Griliches 1991).

[5]The dual approach also allows one to endogenze R&D and capital investment decisions through the use of Euler equations. This is a plus if the specified model is correct but a minus if not.

3 The Model

We use a standard value–added production–function framework similar to that employed by Griliches (1979). In particular, output Q is produced by conventional inputs capital K and labor L. In addition to the conventional inputs, output depends on a stock of knowledge or productivity, which we model as Hicks neutral. Firms' R&D activities augment that stock, which depreciates over time.

Specifically, there are n firms, $i = 1, \ldots, n$, observed in T time periods, $t = 1, \ldots, T$. The production function is

$$Q_{it} = A_{it} M_t U_{it} H_i(x_{it}), \tag{1}$$

where Q is output, $A \times M \times U$ is the state of knowledge or productivity, and x is a vector of conventional inputs. Taking logs we have

$$q_{it} = a_{it} + h_i(x_{it}) + \mu_t + u_{it}, \tag{2}$$

where lower case letters denote logarithms ($a_{it} = \ln A_{it}$, etc.). Productivity consists of two parts: a systematic component a, and a random component $\mu + u$. The random component is further subdivided into an aggregate shock, μ, and an idiosyncratic shock $u \sim F(0, V)$.

We model knowledge acquisition using a standard capital–accumulation type equation. Specifically, firm i's systematic stock of knowledge is

$$S_{it} = (1 - \delta) S_{it-1} + R_{it-1}, \tag{3}$$

where δ is the depreciation rate and R_{it} is firm i's investment in knowledge in period t.[6]

We assume that the systematic component of productivity is a weighted average of the R&D activities of all firms

$$a_{it} = \theta S_{it} + \sum_{j \neq i} w_{ij} S_{jt}, \tag{4}$$

where w_{ij} is a weight that corresponds to some notion of the distance between i and j. Our objective is to uncover the weighting matrix $W = [w_{ij}]$.

We do this as follows. Assume that we have K distance (really proximity) measures, where each measure is an $n \times n$ matrix, $D^k = [d^k_{ij}]$. Let $D_{ij} = (d^1_{ij}, \ldots, d^K_{ij})$ be the vector of measures for the (i, j) pair of firms and define the spillover pool associated with D^k to be

$$S_{-ikt} = \sum_{j \neq i} d^k_{ij} S_{jt}. \tag{5}$$

[6]We follow Hall et al. (2005) in constructing the empirical measure of the R&D stock using a knowledge depreciation rate of 15%. See the data appendix for details.

Based on this definition, the simplest parametric weights correspond to

$$a_{it} = \theta S_{it} + \zeta_k S_{-ikt}, \tag{6}$$

or

$$a_{it} = \theta S_{it} + \sum_k \zeta_k S_{-ikt}. \tag{7}$$

With (6), $w_{ij} = \zeta_k d_{ij}^k$ and with (7), $w_{ij} = \sum_k \zeta_k d_{ij}^k$.

More generally one can incorporate flexibility into the weighting matrix and estimate the associated equation parametrically or nonparametrically. With the latter, each element of W is assumed to be a common function of the vector of distance measures, $w_{ij} = g(D_{ij})$. The vector D must be specified by the practitioner. However, the functional form of g is determined by the data.

4 Measuring Spillovers

Alternative distance or proximity measures have been used in the literature. We build on and refine some of these.

Technology Spillovers

We, like most researchers who examine technological proximity, use a measure that is due to Jaffe (1986). Suppose that there are L technology clases, $l = 1, \ldots, L$. Let F_{il}^T be the fraction of firm i's R&D activities (e.g., expenditures or patents) that are in class l. i's technology locational distribution is then $F_i^T = (F_{i1}^T, F_{i2}^T, \ldots, F_{iL}^T)$. The measure of technology match between firms i and j is the uncentered correlation coefficient

$$w_{ij}^T = \frac{\sum_{l=1}^L F_{il}^T F_{jl}^T}{\sqrt{[\sum_{l=1}^L (F_{il}^T)^2][\sum_{l=1}^L (F_{jl}^T)^2]}}. \tag{8}$$

Notice that with (8) all classes are treated symmetrically — no class is 'closer' to any other. Furthermore, spillovers occur within but not across classes. In other words, there are no $F_{il}^T F_{jk}^T$ terms with $l \neq k$.

Horizontal Spillovers

Most people use a discrete measure, same or different Standard Industrial Classification (SIC), that distinguishes between intra and interindustry spillovers (e.g., Levin & Reis 1988, Bernstein 1988, Bernstein & Nadiri 1989, Ornaghi 2006). We use a continuous and more disaggregate measure that was employed by Bloom, Shankerman, & Van Reenen (2008). In particular, suppose that there are M product markets, $m = 1, \ldots, M$, where markets are proxied by SICs, and let F^P be defined similarly to F^T, i.e., F^P_{im} is the fraction of i's output that belongs to product market m. Then F^P_i is firm i's product–market locational distribution, and w^P_{ij} is the uncentered correlation coefficient between the output distributions F^P_i and F^P_j. Notice that here, as with the technology measure, product markets are treated symmetrically — no market is 'closer' to any other — and spillovers occur within but not across markets.[7]

Geographic Spillovers

Most previous studies focus on the location of firms' headquarters as the basis for measuring geographic spillovers. Several measures have been used: i) the Euclidean distance between the capital cities of the countries where the headquarters of firms i and j are located (Eaton & Kortum 1996, Keller 2002) and ii) a 0/1 variable that distinguishes if the firms' headquarters are located in a different or the same region (Eaton & Kortum 1999, Orlando 2004).

It is not clear, however, that headquarters is the relevant locational variable. It seems to us more likely that inventors communicate, and many inventors are employed in research labs. Moreover, unlike headquarters, a given firm might locate laboratories in many regions. Although the locations of headquarters and labs are correlated, correlation is far from perfect. To illustrate, Microsoft began operations in the Seattle area, where it located both its head office and its research lab. However, over time, it constructed additional labs in high–tech areas of the US, such as Silicon Valley and the Boston area. Still more recently, it has developed international labs in places as distant as Bangalore and Beijing.

In our empirical work we look at the locations of both headquarters and research labs. However, we do not have data on the geographic locations of labs. To circumvent this problem, we use patent data. Specifically, each patent gives the address of its principal inventor, and we construct geographic R&D locational distributions from those addresses.

We do this as follows. Suppose that there are K geographic regions, $k = 1, \ldots, K$. Let F^G_{ik} be the fraction of firm i's inventors that are located in region k, and F^G_i be firm i's geographic–market locational distribution. All of our geographic measures are of the form

$$w^G_{ij} = \sum_k \sum_l f(F^G_{ik}, F^G_{il}, F^G_{jk}, F^G_{jl}) C(d_{kl}), \quad i \neq j, \quad w^G_{ii} = 0, \tag{9}$$

where d_{kl} is the Euclidean distance between regions k and l.

[7]The revised version of Bloom et al (2008) considers extending this to measures which do take into account the idea that some industries (and technological classes) may be closer to each other by using the Mahalanobis distance metric instead of the Jaffe style measure.

We distinguish between measures of *match* (f) and measures of *Euclidean distance* (C).

i) Measures of match

We use two functions $f(.)$ of the locational distributions:

The *product* function,

$$f_{\text{prod}}(.) = F^G_{ik} F^G_{jl}, \tag{10}$$

and the *sum* function

$$f_{\text{sum}}(.) = O_{ijkl}(F^G_{ik} + F^G_{il} + F^G_{jk} + F^G_{jl}), \tag{11}$$

where O_{ijkl} equals 1 if there is an overlap between the activities of firms i and j in the combined regions k and l. The product function is a multidimensional relative of the uncentered correlation coefficient, whereas the sum function measures, for those regions where the two firms' activities overlap, the fraction of their combined activities that are located in the combined regions.

There are a number of attributes of our geographic measures that should be noted. First, unlike our product market and technology class measures, with our geographic measures spillovers occur across markets and cross–market spillovers are weighted by geographic distance between the markets. Note also that the second measure (the sum) favors diversification of labs. In other words, if a rival's inventors are distributed across all locations, firm i would also like to diversify its locations. In contrast, with the product measure, firm i is indifferent between specialization and diversification. Finally, note that each measure specializes to headquarters, which is similar to considering firms that have only one lab or have all of their labs in the same region.

ii) Measures of Euclidean distance

With our parametric specifications, we also use alternative functions of d_{kl}. Specifically, we define two simple functions $C(.)$ of the Euclidean distance between regions k and l.

$$C_{\text{corr}}(d_{kl}) = \begin{cases} 1, & k = l, \\ 0, & k \neq l, \end{cases} \tag{12}$$

and

$$C_{\text{exp}}(d_{kl}) = \exp(-\alpha d_{kl}). \tag{13}$$

With the first function, geographic distance does not matter. Indeed, no spillovers occur unless firms i and j are located in the same region. We use the subscript *corr* to indicate that this is the distance measure that is implicit when the uncentered correlation coefficient is calculated (i.e., our measures of technology and product–market spillovers). With the second measure, the effect of distance decays gradually. The parameter α in (13), which determines the rate of decay, can be estimated or set exogenously. We use the subscript *exp* to denote exponential decay.

5 Data

We use two firm level data sources. Accounting data (sales, employment, capital, etc.) come from U.S. Compustat 1980-2000. Since we focus on manufacturing, we removed all firms whose primary industries are outside of the manufacturing sector (SIC codes 2000-3990). The data items available in Compustat, which are those reported in standard corporate income statements, are used to construct our measures of output and conventional inputs, Q, L, and K. We deflate sales by a year–specific four digit producer price index and capital by the CPI. The Compustat Segment data also allow us to create breakdowns of each firm's activities across product markets. Specifically, firms' sales are allocated to industry codes or lines of business, such as 3660 ("Communications equipment") and 3820 ("Measuring and controlling devices"). On average each firm reports sales in approximately 4.8 different industry codes.

The Compustat firms were matched into the U.S. Patent and Trademark Office (USPTO) data from the NBER data archive. This archive contains detailed information on almost three million U.S. patents granted between January 1963 and December 1999 and all citations made to those patents between 1975 and 1999.[8] Since our method requires information on patenting, we kept only those firm years with a positive patent stock (firms that had no patents at all in the period of 1970-2000 were dropped), leaving an unbalanced panel of 812 manufacturing firms with at least four observations with non-missing sales, capital stock and employment information. The USPTO allocates patents into technology classes, such as class 042 ("Firearms") and class 257 ("Solid state devices"), the breakdown that we use to construct technological proximity. On average each firm owns 496 patents in 37 classes.

Inventor location is taken from the address of the lead inventor of the patent, which is recorded at the city level. This is the standard measure of inventor location used by Jaffe and Trajtenberg (2002) and Griffith , Harrison and Van Reenen (2006), among many others. We feel that it is a more appropriate proxy for the location of the firm's R&D, and thus the potential for spillovers, than the headquarters of the firm, as it is a better indicator of where the key research was conducted. We allocate R&D activity into 2,039 geographic units, where a unit is a county. Because there are multiple patents, we are able to build

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

Table 1: Descriptive Statistics

Variable	(1) Mean	(2) Median	(3) Std.dev.
Sales, Q	2,832	349	10,320
Stock of physical capital, K	1094	94	4,203
Employment, L	13,582	2,600	43,631
R&D stock, S	546	47	2,446
Technological spillovers, SpillTech	30,670	26,034	21,296
Product market spillovers, SpillSIC	35,855	20,740	39,393
Geographic market spillovers, SpillGeog	42,790	42,101	19,227

Notes:
The means, medians and standard deviations are taken over all non-missing observations between 1980 and 2000; values measured in 2000 prices in $million
The geographic spillover measure is based on inventor locations and uses the product measure of match and exponential decay with distance

up a picture of the location of the firm's R&D activity spatially. We do not use inventor information outside the United States because we have focused on US firms.

Each firm's own stock of knowledge, S, is constructed from R&D expenditure data as in equation (3) with the depreciation rate δ set to 0.15 following Hall, Jaffe, and Trajtenberg (2005). The horizontal product–market spillover measure (with weights w^P), which is denoted SpillSIC, makes use of 349 3 digit SICs, and the technological spillover measure (with weights w^T), which is denoted SpillTech, makes use of 410 technology classes. Finally, the measures of geographic proximity (with weights w^G), which are denoted SpillGeog, differ depending on the functions of match (f) and geographic distance (C) that are chosen.

5.1 Descriptive Statistics

Table 1 provides some basic descriptive statistics for the accounting and patenting data, and the technology, product market, and geographic distance measures, SpillTech, SpillSIC, and SpillGeog. The sample firms are large (mean employment is over 13,000), but with much heterogeneity in size (measured by sales, employment, or physical capital), R&D intensity, patenting activity, and location. The three spillover measures also differ widely across firms.

In order to distinguish between the effects of technology, product market, and geographic spillovers econometrically we need independent variation in the distance metrics in technology, product market, and geographic space. To gauge this we do two things.

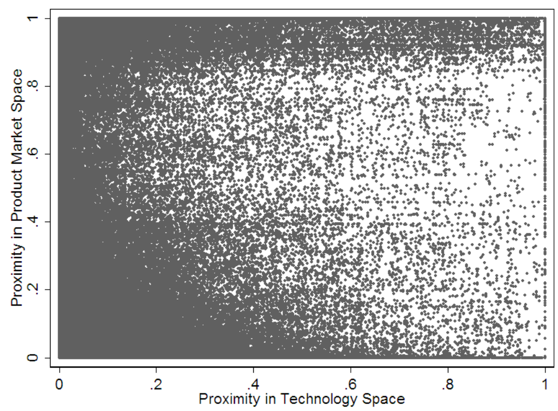
First, we calculate and report in table 2 raw correlation coefficients between pairs of

Table 2: Correlations between the spillover measures

Correlation between	(1) SIC–Tech	(2) SIC–Geog	(3) Geog–Tech
weights, w	0.359	0.040	0.038
lnSpillM	0.420	0.162	0.220
the change of lnSpillM	0.460	0.129	0.327

Notes:
All correlations are significant at the 1% level
In column one, M denotes SIC, TECH, or Geog
The change in lnSpill is defined as a residual from the regression with firm fixed effects and time dummies.

Figure 1: SIC and TECH correlations



measures. To interpret table 2, consider the first column. The first row in that column contains the correlation coefficient between the product and technology market weights, w^P and w^T, the second column that between lnSpillSIC and lnSpillTech, and the third between ΔlnSpillSIC and ΔlnSpillTech. In performing this calculation, we used the product measure of match and the exponential measure of distance to construct SpillGeog. The correlations between the weights, the levels, and the changes of the spillover variables are positive and significant at the 1% level. However, they are well below unity, implying substantial independent variation in each of the three measures.

Second, we plot each pair of weights. Figures 1, 2, and 3, which show w^T, w^P, and w^G plotted against each other reveal that the positive correlations that we observe are caused by a dispersion across the unit box rather than by a few outliers.

We also investigate the relationship between the locations of inventors and headquarters in two ways. First, we calculate the aggregate distribution of inventor locations, which shows the aggregate fraction of inventors that are located at various distances from their

Figure 2: SIC and GEOG correlations

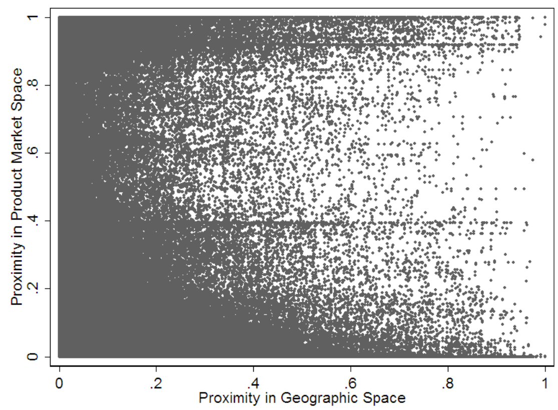


Figure 3: GEOG and TECH correlations

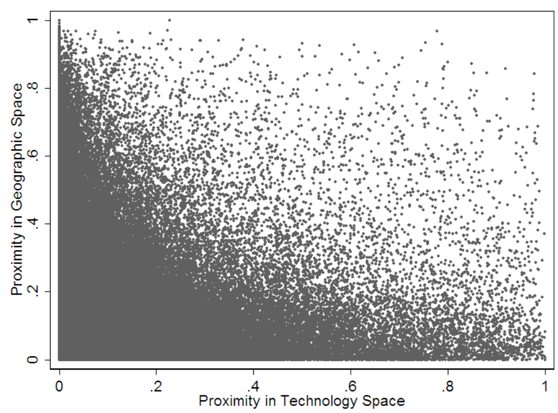


Table 3: Distance between inventors and headquarters

Percentile	0%	20%	40%	60%	80%	100%
Distance	0 km	0 km	52 km	414 km	1656 km	8097 km

headquarters. The quintiles of this distribution are reported in table 3, which shows that although some firms have all of their research activities located in close proximity to their headquarters, the activities of others are much more dispersed.

Second, we picked four firms that exhibit very different patterns. The information for those firms is summarized in figure 4. The triangles in the graphs locate each firm's headquarters, whereas the circles show the distribution of its inventors. The first firm, Polaroid, has almost all of its research activities located in close proximity to its headquarters; the second, Motorola, has four concentrations of inventors, one of which is in close to its headquarters; the third, Eaton (a power management company), has many concentrations of research activities. This dispersion is due to the fact that Eaton has expanded mainly through acquisitions; finally, the fourth, Union Carbide, has most of its R&D activities located in the Northeast where its headquarters used to be. In 1982, however, it relocated its headquarters to Connecticut and later to Texas but left its labs in place.

6 Econometric Methodology

6.1 The Model

We use a Cobb–Douglas production function. The dependent variable is lnQ, where Q is real sales, and the explanatory variables are lnK, where K is capital, and lnL, where L is labor, as well as firm and year fixed effects. In addition, we include a firm's own stock of knowledge, lnOwnS, the horizontal product–market spillover measure, lnSpillSIC, and the technological spillover measure, lnSpillTech. Most equations that are reported below contain all of these variables. In addition, they contain one or more measures of geographic proximity, lnSpillGeog, that differ depending on the functions of match (f) and geographic distance (C) that are chosen. Variables that are preceded by ln are in natural logarithms.

In some cases, our construction of the geographic spillover measure is somewhat different from that of the product–market and technology spillover measures, lnSpillSIC and lnSpillTech. We define two geographic spillover variables,

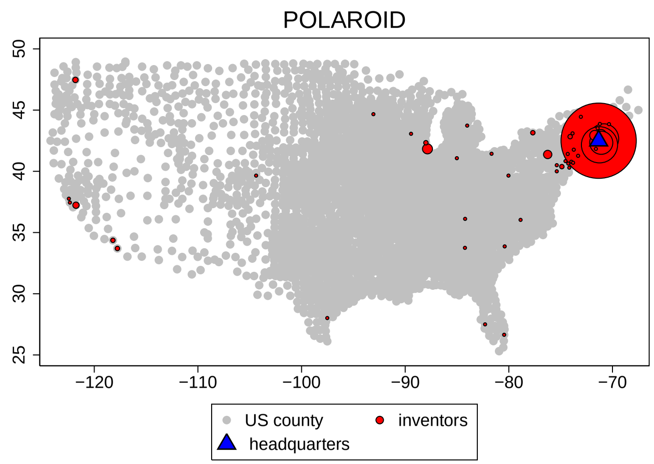
$$\text{lnSpillGeog}_{it} = \ln\Bigl(\sum_j w_{ij}^G S_{jt}\Bigr), \tag{14}$$

and

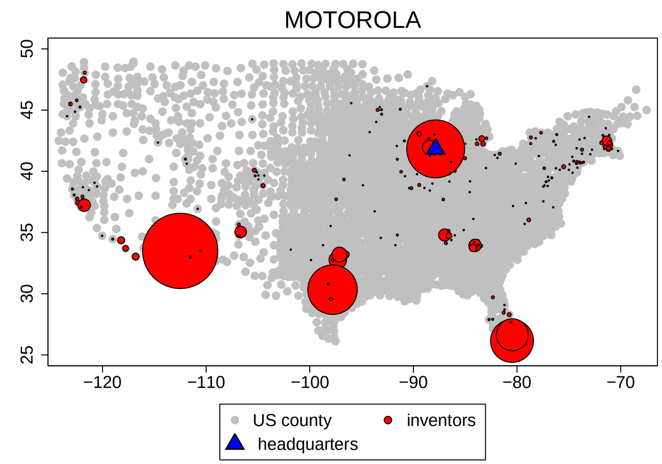
$$\text{SpillGeogln}_{it} = \sum_j w_{ij}^G \ln(1 + S_{jt}). \tag{15}$$

We make this distinction because using the logarithm of sums makes semiparametric estimation difficult. As one substitutes a series approximation for w_{ij}^G, the production–function equation becomes nonlinear in the unknown parameters, and the estimates can only be obtained from numerical minimization. The problem is that the objective function is not convex; it has many local minima, their number growing with the number of terms in the series expansion. Using the sum of logarithms, in contrast, makes the equation linear in parameters, and one can estimate it simply by applying OLS.

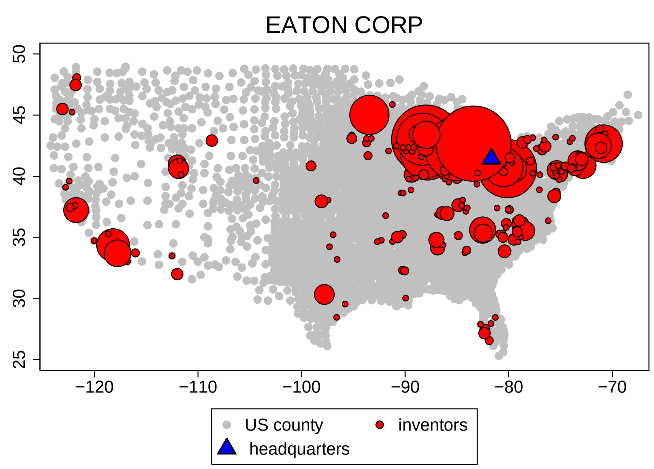
We therefore faced a tradeoff between being confident that the global minimum was found, which is difficult to ensure with (15), and being consistent in the definition of the spillover measure. We chose the former (log–of–sums) for the parametric exercise. All



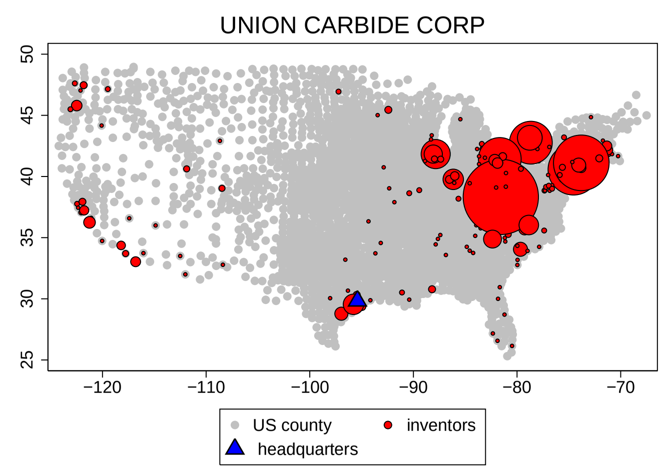
Polaroid – inventors are concentrated around the headquarters



Motorola – four clusters in Illinois, Texas, Arizona, Florida. According to the company's website, these are the locations of R&D centers



Eaton Corp — a very diversified company — is involved in joint ventures with other big businesses, expanded mainly through acquisitions. As a result, the inventors are scattered all over the Northeastern region



Union Carbide – moved its headquarters to Connecticut in 1982, and then to Texas. R&D activity remained in New York, New Jersey and West Virginia

Figure 4: Inventor Distributions

semiparametric estimates reported, however, use the sum–of–logs form.

Our estimating equation is thus

$$y_{it} = \beta_0 + X_{it}\beta + \nu_i + \mu_t + u_{it}, \qquad i = 1, \ldots, n_t, \qquad t = 1, \ldots, T_i, \tag{16}$$

with $y_{it} = \ln\mathrm{Q}_{it}$ and

$$X_{it} = [\ln\mathrm{K}_{it}, \ln\mathrm{L}_{it}, \ln\mathrm{OwnS}_{it}, \ln\mathrm{SpillSIC}_{it}, \ln\mathrm{SpillTech}_{it}, \ln\mathrm{SpillGeog}_{it}], \tag{17}$$

where, in the semiparametric case, lnSpillGeog is replaced with SpillGeogln. Note that we have a doubly unbalanced panel, since n differs by year and T differs by firm.

6.2 Econometric Issues

i) Identification

The potential bias in OLS estimates of production functions has long been recognized (see, e.g., Marschak and Andrews 1944). This bias results from the possible correlation between input levels and firm level productivity shocks. Specifically, when firms experience a large productivity shock, they might respond by using more inputs. Applied economists have devised alternatives to OLS that circumvent this problem. Most use either a variant of the method developed by Olley and Pakes (1996) and extended by Levinsohn and Petrin (2003) or the GMM methods proposed by Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (2000). We have chosen to focus on GMM approaches because it is difficult to introduce endogenous R&D into Olley–Pakes approaches.

Essentially, the Arellano and Bond approach assumes that serial correlation of the error term (after controling for fixed effects) is of finite order. In the simplest case of no serial correlation this implies that levels dated $t-2$ and earlier are valid instruments. If there is some first order correlation (e.g. MA(1)) we can use $t-3$ dated instruments. We also considered the additional moments for the levels equations suggested by Arellano and Bover (1995) and first used in a production function context by Blundell and Bond (2000).

We assume that the inventor distances are exogenous and instrument for the terms in the semiparametric expansion by the corresponding terms with potentially endogenous regressors replaced with corresponding instruments (see Pinkse, Slade, and Brett, 2002). The assumption that inventor distances are exogenous is questionable. However, in our model inventor locations are constant over the observed time period. Moreover, finding plausible instruments for such distances is near impossible in the present application, unlike in Pinkse, Slade, and Brett (2002).

ii) Estimation

We use four different methods to estimate the unknown coefficients in (16): OLS, GMM, semiparametric least squares (SLS) and semiparametric GMM (SGMM). For (S)GMM we consider both a static and a dynamic version.

The parametric methods presume that we know the shape of the weights in lnSpillGeog, which is likely unreasonable. As is often the case, our choice is a tradeoff between 'bias' arising from using a parsimonious model and increased 'variance' arising from having too many unknowns.

For the static models we subtract out firm–specific means in (16) to eliminate firm fixed effects. We then obtain

$$y_i^* = X_i^* \beta + \mu^* + u_i^*, \tag{18}$$

where $y_i^*, X_i^*, \mu^*, u_i^*$ are demeaned versions of y_i, X_i, μ, u_i respectively. With OLS and SLS there is an implicit and impausible assumption of strict exogeneity. Errors here can be serially correlated.

For the static versions of GMM and SGMM, we use sufficiently lagged x_{it}'s as instruments. This presumes that such instruments are orthogonal to current and (one period–) lagged errors. Such orthogonality of instruments and errors for all intents and purposes entails either strict exogeneity of the instruments or an absence of serial correlation in the errors. Our choice of using lagged x_{it}'s as instruments implies the second interpretation.

For the dynamic version of the model we use the methodology of Arellano and Bond (1991) and estimate a model similar to that of Blundell and Bond (2000). Consider equation (16), but allow the error term to follow the AR(1) process

$$u_{it} = \rho u_{it-1} + \varepsilon_{it},$$

where ε_{it} is serially independent (we can allow this to be correlated up to finite order). Substituting this back into equation (18) gives

$$y_{it} = \pi_1 y_{it-1} + X_{it}\pi_2 + X_{it-1}\pi_3 + \nu_i^* + \mu_t^* + \varepsilon_{it}. \tag{19}$$

The production function estimates are identified using the common factor restriction:

$$\beta = \pi_2 = -\frac{\pi_3}{\pi_1}. \tag{20}$$

We estimate the unrestricted model of equation (19) and then impose (20) by minimum distance methods.

iii) Semiparametric Estimation

For our semiparametric estimation of the distance function C, as before, the estimated equation is (18). However, as discussed in subsection 6.1, lnSpillGeog is replaced by SpillGeogln.

We follow Chen (2007) (see also Dechevsky and Penev, 1997) and approximate C by

$$C(d) = \sum_{m=0}^{2^{J_n}} \gamma_m b\big(2^{J_n}\tilde{G}(d) - m + 1\big),$$

where

$$b(x) = \begin{cases} 1 - |x-1|, & |x-1| < 2, \\ 0, & |x-1| \geq 1, \end{cases}$$

and $\tilde{G}$ is a monotonic distance transformation described below.

The advantage of our choice of basis functions is that it allows us to control the shape of C by imposing restrictions on the γ–coefficients. This is important to ensure that C decrease to zero sufficiently fast; see Pinkse and Slade (2010) for further discussion.

As with all nonparametric methods there is a sample–size–dependent input parameter, here J_n. J_n should increase with the sample size n to an infinite limit. In finite size samples, however, this requirement provides little guidance, and we experiment with two different values for J_n: $J_n = 2$ (5 basis functions) and $J_n = 3$ (9 basis functions).

In some specifications, we impose that our γ–coefficients satisfy the restrictions $\gamma_{2^{J_n}} = 0$ and

$$\forall 0 \leq \tilde{m} \leq m \leq 2^{J_n} : \gamma_m \leq \gamma_{\tilde{m}} \exp\Big[-\alpha\Big(\tilde{G}^{-1}(m/2^{J_n}) - \tilde{G}^{-1}(\tilde{m}/2^{J_n})\Big)\Big], \tag{21}$$

for $\alpha = 0$ and $\alpha = (\ln 2)/3000$, which results in monotonically decreasing C and both exponentially and monotonically decreasing C, respectively.

Finally, if d is distance in thousands of kilometers and $[d]$ is the largest integer no greater than d, then $\tilde{G}(d) = (d - [d])\tilde{G}_{[d]+1} + \big(1 - (d - [d])\big)\tilde{G}_{[d]}$, where $\tilde{G}_i$ is the fraction of inventor pairs at different firms who are no more than $1000i$ kilometers apart.

For SGMM we again apply the Arellano and Bond (1991) technique, albeit that we now need a sample–size dependent number of instruments. We obtain such instruments using a simplified version of the estimation procedure of Pinkse, Slade and Brett (2002).

iv) Inference

For the parametric result, we use fairly standard methodology, albeit that we make some independence assumptions across firms for the errors. More precisely, we assume the existence of constants σ_{ti} such that

$$E[u_{it}u_{js}|z_{it}, z_{js}] = \begin{cases} 0, & j \neq i, \\ \sigma_{|t-s|,i}, & j = i. \end{cases} \tag{22}$$

For most specifications, we set $\sigma_{|t-s|,i} = 0$ if $|t-s| \geq r$, where r is a parameter that we choose. The difference between (22) and full 'clustering' is that the covariance between u_{it} and u_{is} depends only on the identity of the firm and the difference $|t-s|$ and that the conditional covariance of u_{it}, u_{is} does not depend on z_{it}, z_{is}. Our choice is motivated by the idea of providing a reasonable compromise between the extremes of Stock and Watson (2008) and full clustering.

Asymptotics for semiparametric IV estimators under spatial dependence are derived by Pinkse, Slade, and Brett (2002), who establish consistency and asymptotic normality. However, their results do not allow for the presence of inequality restrictions like those imposed to achieve monotonicity and exponential decline, they do not explicitly address the issue of Comfac restrictions,[9] and the method of generating spillover measures based on inventor locations complicates matters further.

An earlier version of this work contained standard errors and confidence bands based on the Pinkse, Slade, and Brett (2002) method, again using a compromise between the extremes of full clustering and Stock and Watson (2008). The results presented here are based on *subsampling*[10] and yield qualitatively the same results. We opted for subsampling since the requirements for its consistency are the weakest among possible alternatives. Nevertheless, we do not rigorously establish that our estimator has a limiting distribution nor indeed a convergence rate, and some of the issues mentioned in the first paragraph are problematic with subsampling, also. So standard errors and confidence bands for the semiparametric case should be viewed with some caution, but the qualitative similarity of results based on the two methods mentioned is reassuring.

7 Results

The remaining tables contain a number of specifications of our production function with knowledge spillovers, all of which include firm and year fixed effects. We begin with the OLS estimates.

[9]Comfac would be a comparatively straightforward adjustment.

[10]Subsampling is a resampling method akin to the bootstrap. It requires weaker conditions than the bootstrap at the cost of slower convergence if the bootstrap conditions are satisfied; see Politis, Romano, and Wolf (1999).

7.1 Fixed Effects OLS

Single Spillover Measures

Table 4, which assesses each spillover measure in isolation as well as all three jointly, shows that when only one measure is considered, the coefficient of that measure is positive and significant at conventional levels. However, the coefficients of lnSpillTech and lnSpillGeog are an order of magnitude larger than the coefficient of lnSpillSIC. Furthermore, when all three measures are included in the same equation, the coefficient of lnSpillSIC becomes negative and insignificant. We will see that the negative insignificant coefficient persists in virtually all subsequent specifications.

This is consistent with the model outlined in Bloom et al (2008) where R&D by product market rivals affects market value, but should have no effect on total factor productivity (TFP), as it does not alter technological capabilities (conditional on own R&D). If we have not properly controlled for firm-specific prices when measuring output, however, then SpillSIC could have a negative effect in the (revenue-based) production function as it depresses own prices.

Multiple Spillover Measures

Tables 5–7 contain further estimates of the production function. The geographic measures in table 5 are based on inventor locations, those in table 6 are based on headquarters, whereas table 7 assesses inventors and headquarters based measures jointly. Looking across equations and tables, we can see that, in all specifications, the coefficients of the common variables are remarkably similar in magnitude and significance. Furthermore, all of those coefficients are positive, significant at high levels, and the capital and labor coefficients are similar in magnitude to those obtained by previous researchers.[11] Finally, as in Bloom, Schankerman, and Van Reenen (2008), the coefficient of the technology spillover variable is substantially larger than that of the horizontal measure.

To examine geographic spillovers in more detail, first consider table 5, which is based on inventor locations and contains all combinations of our measures of match (*sum* and *prod*) and distance (*corr* and *exp*). When considered separately, the estimates of the coefficients of the *sum* measure are larger than those of the *prod* measure. However, when both are included in the same equation, the coefficients of the *prod* measure become larger than those of the *sum* measure. Moreover, with the exponential distance function, the latter become negative and insignificant whereas the former remain virtually unchanged. Finally, comparing across specifications, we see that the *sum* coefficients are highly unstable. For

[11] Note that our estimates show mildly diminishing returns when only the firm's choice variables are considered but strongly increasing returns when spillovers are added.

Table 4: OLS Production Function Estimates, Single Proximity Measures

Dependent variable is ln(sales)								
	(1)		(2)		(3)		(4)	
	SpillSIC		SpillTech		SpillGeog Prod,Exp		All Three Prod,Exp	
lnK	0.249	(0.008)	0.234	(0.008)	0.234	(0.008)	0.225	(0.008)
lnL	0.599	(0.010)	0.606	(0.010)	0.610	(0.010)	0.614	(0.010)
lnOwnS	0.077	(0.005)	0.060	(0.005)	0.066	(0.005)	0.054	(0.005)
lnSpillSIC	0.073	(0.017)					-0.025	(0.018)
lnSpillTech			0.550	(0.035)			0.395	(0.039)
lnSpillGeog					0.607	(0.028)	0.524	(0.029)

Notes:
Inventor location measures
Firm and year fixed effects included.
Standard errors are in parentheses (homoskedasticity is assumed)
Distance weight declines at a rate of 50% per 200km

Table 5: OLS Production Function Estimates, Inventors

Dependent variable is ln(sales)						
	(1)	(2)	(3)	(4)	(5)	(6)
	Sum, Corr	Prod, Corr	Both, Corr	Sum, Exp	Prod, Exp	Both, Exp
lnK	0.234	0.230	0.231	0.234	0.225	0.225
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
lnL	0.609	0.611	0.611	0.611	0.614	0.613
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
lnOwnS	0.059	0.060	0.059	0.055	0.054	0.054
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
lnSpillSIC	-0.038	-0.029	-0.030	-0.038	-0.025	-0.025
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
lnSpillTech	0.524	0.490	0.485	0.468	0.395	0.400
	(0.039)	(0.040)	(0.040)	(0.040)	(0.039)	(0.040)
lnSpillGeog(sum)	0.464		0.169	0.920		-0.104
	(0.066)		(0.083)	(0.085)		(0.110)
lnSpillGeog(prod)		0.281	0.232		0.524	0.548
		(0.031)	(0.040)		(0.029)	(0.039)

Notes:
lnSpillGeog is standardized: mean = 0, standard deviation = 1.
Firm and year fixed effects included.
Standard errors are in parentheses (homoskedasticity is assumed)
Exponential weight declines at a rate of 50% per 200km

Table 6: OLS Production Function Estimates, Headquarters

Dependent variable is ln(sales)				
	(1)		(2)	
	Prod, Corr		Prod,Exp	
lnK	0.235	(0.008)	0.229	(0.008)
lnL	0.606	(0.010)	0.606	(0.010)
lnOwnS	0.061	(0.005)	0.055	(0.005)
lnSpillSIC	-0.047	(0.018)	-0.045	(0.018)
lnSpillTech	0.564	(0.039)	0.528	(0.039)
lnSpillGeog(prod)	0.087	(0.014)	0.298	(0.022)

Notes:
lnSpillGeog standardized: mean 0, standard deviation 1
Firm and year fixed effects included.
Standard errors are in parentheses (homoskedasticity is assumed)
Exponential weight declines at a rate of 50% per 200km

Table 7: OLS Production Function Estimates, Headquarters and Inventors

Dependent variable is ln(sales)				
	(1)		(2)	
	Prod, Corr		Prod,Exp	
lnK	0.232	(0.008)	0.225	(0.008)
lnL	0.610	(0.010)	0.613	(0.010)
lnOwnS	0.059	(0.005)	0.053	(0.005)
lnSpillSIC	-0.036	(0.018)	-0.028	(0.018)
lnSpillTech	0.489	(0.040)	0.403	(0.040)
lnSpillGeog(inventors)	0.246	(0.032)	0.476	(0.040)
lnSpillGeog(HQ)	0.062	(0.014)	0.053	(0.030)

Notes:
lnSpillGeog is standardized: mean = 0, standard deviation = 1.
Firm and year fixed effects included.
Standard errors are in parentheses (homoskedasticity is assumed)
Exponential weight declines at a rate of 50% per 200km

Table 8: Production Function Estimates and t–statistics, Covariance Estimation

Dependent variable is ln(sales)							
	(1) Coef.	(2) Homo. t stat	(3) Hetero. t stat	(4) Cluster. $r = 0$	(5) Cluster. $r = 2$	(6) Cluster. $r = 21$	(7) Cluster. Unrestr.
lnK	0.225	29	21.5	24.1	14.5	11.9	11.1
lnL	0.614	62.9	45.4	52.1	31.8	25.4	23.6
lnOwnS	0.054	10.2	8.17	9.38	5.42	4.09	4.07
lnSpillSIC	-0.025	-1.42	-1.4	-1.58	-0.88	-0.65	-0.64
lnSpillTech	0.395	10	8.07	9.22	4.85	3.1	3.11
lnSpillGeog	0.524	17.9	11	13.4	6.86	4.16	4.15

Notes:
Inventor geographic measures.
Measure of match: product. Distance weight: exponential (half-life distance 200km)
lnSpillGeog is standardized: mean = 0, standard deviation = 1.
Firm and year fixed effects included.
Columns (2) through (7) contain t–statistics for the indicated choices of covariance matrix estimate, where in columns (4) through (6) dependence up to lag no greater than r is assumed.
Columns (3) and (7) allow for heteroskedasticity across time and firm, whereas columns (4) through (6) allow for heteroskedasticity across firms, but not time.

these reasons, in what follows, we report only equations that use the product measure.

Table 5 also contains specifications that use the *corr* and *exp* functions of geographic distance. It is clear that, although the coefficients of both measures are positive and significant at conventional levels, with one exception the magnitude of the spillover effect is much larger when the effect of distance is allowed to decay gradually. In other words, a zero/one distinction (different or same geographic market) has much less explanatory power, implying that cross–market spillovers are important.

Table 6, which contains measures based on headquarters, shows similar patterns. In particular, the estimate of the coefficient of the spillover measure that uses the *corr* distance function is considerably smaller than that of the measure that uses the *exp* function. Finally table 7, which assesses measures based on inventor and headquarters locations jointly, shows that inventor measures outperform headquarters measures.

Tables 5–7 thus reveal four regularities that persist in alternative specifications. First, the *prod* measure of match, which is the natural extension of the uncentered correlation coefficient, outperforms the *sum* measure. Second, although match by itself is important, it is clear that considering both match and Euclidean distance is associated with estimates of spillover effects that are both larger and more significant than when distance is ignored. Third, geographic distance effects are larger when inventor rather than headquarter locations are chosen. Our use of the patent data therefore appears to result in an improved geographic measure. Finally, the coefficients of the spillover variables are substantially larger that those of the own stock of knowledge, another finding that is consistent with previous research.

Estimates in tables 5–7 use standardized geographic spillover variables.[12] Standardization makes it more difficult to interpret the values of coefficients. To demonstrate the magnitude of geographic spillovers we perform the following thought experiment.[13] We track the total factor productivity (TFP) of a hypothetical firm, whose inventors are located in Los Angeles County ("incumbent LA firm"). We assume that another hypothetical firm has a lab in the same county with an R&D stock of $500m. We then relocate this lab farther and farther away from LA. As the lab moves to an infinite distance, its contribution to the LA incumbent's geographic spillover variable declines to zero, while the product market and the technology spillovers are left unaffected. Having such a lab in LA raises the incumbent LA firm's TFP by 0.82%, whereas having the mobile R&D lab in San Diego would raise TFP by only 0.44%. If the lab were moved to San Francisco, its contribution would decline further to 0.12%. Having the lab as far as Seattle would only raise the LA–based firm's TFP by 0.00038% (essentially zero). Since LA's GDP was approximately $55 billion in 2002 according to the Census, this implies the lab with a $500m R&D stock raises LA county's net output by $451 million. This is a substantial effect and could rationalise why cities and states compete so eagerly for the location of R&D activities (e.g. Wilson, 2009).

These OLS estimates, which suffer from a number of econometric problems that are discussed in section 6, are only preliminary. Nevertheless, they serve as a guide to our more rigorous specifications. In particular, in what follows, we report specifications based on inventor locations, product measures of match, and nontrivial functions of Euclidean distance (C_{exp} or a semiparametric specification).

7.2 Corrected Standard Errors

Table 8 contains t statistics for six different estimates of the covariance matrix. All are based on inventor locations, the f_{prod} measure of match, and the C_{exp} function of distance. The first equation was obtained under the assumption of homoskedasticity, whereas the second assumes unrestricted heteroskedasticity. Neither allows for serial correlation. The remaining entries in the table are based on the assumption that errors cluster by firm but make different assumptions concerning the nature of serial correlation. In particular, the third, fourth, and fifth equations, which allow for restricted correlation, correspond to r

[12]See appendix A for an explanation.

[13] For this experiment, we reestimated the equation in table 5 column (5) without standardizing the geographic spillover variable.

= 0, r = 2, and r = 21, where serial correlation dies out after r periods. Finally, the last equation allows for unrestricted serial correlation.

The table shows that the t statistics tend to decline as we move from left to right. This occurs because the number of coefficients that must be estimated grows. The single exception is the move from specification two to three, since clustering without serial correlation involves fewer parameters than unrestricted heteroskedasticity.

Two regularities are notable. First, the coefficients of the geographic and technology spillover measures are positive and significant, regardless of covariance–matrix estimator. Second, the coefficient of the product–market measure is negative and insignificant in all specifications

Although we do not report further estimates, the qualitative nature of the regularities than we have noted persists when other measures of geographic match and distance are used.

7.3 Semiparametric Estimates

Table 9: Semiparametric estimates, parametric part

Dependent variable is ln(sales)						
	$J_n = 2$			$J_n = 3$		
	(1) no constr.	(2) monot.	(3) exp. decl.	(4) no constr.	(5) monoton.	(6) exp. decl.
lnK	0.231^{**}	0.232^{**}	0.232^{**}	0.226^{**}	0.229^{**}	0.230^{**}
lnL	0.614^{**}	0.612^{**}	0.612^{**}	0.611^{**}	0.613^{**}	0.614^{**}
lnOwnS	0.052^{**}	0.057^{**}	0.058^{**}	0.053^{**}	0.053^{**}	0.054^{**}
lnSpillSIC	-0.044	-0.040	-0.038	-0.024	-0.034	-0.035
lnSpillTech	0.452^{**}	0.505^{**}	0.497^{**}	0.434^{**}	0.441^{**}	0.428^{**}

Notes:
Firm and year fixed effects included.
Inventor geographic measures, product measure of match.
Significance levels are obtained by subsampling
Number of subsamples: 1,000. Size of the subsample: 200 firms
Significance level: † – 10%, * – 5%, ** – 1%.

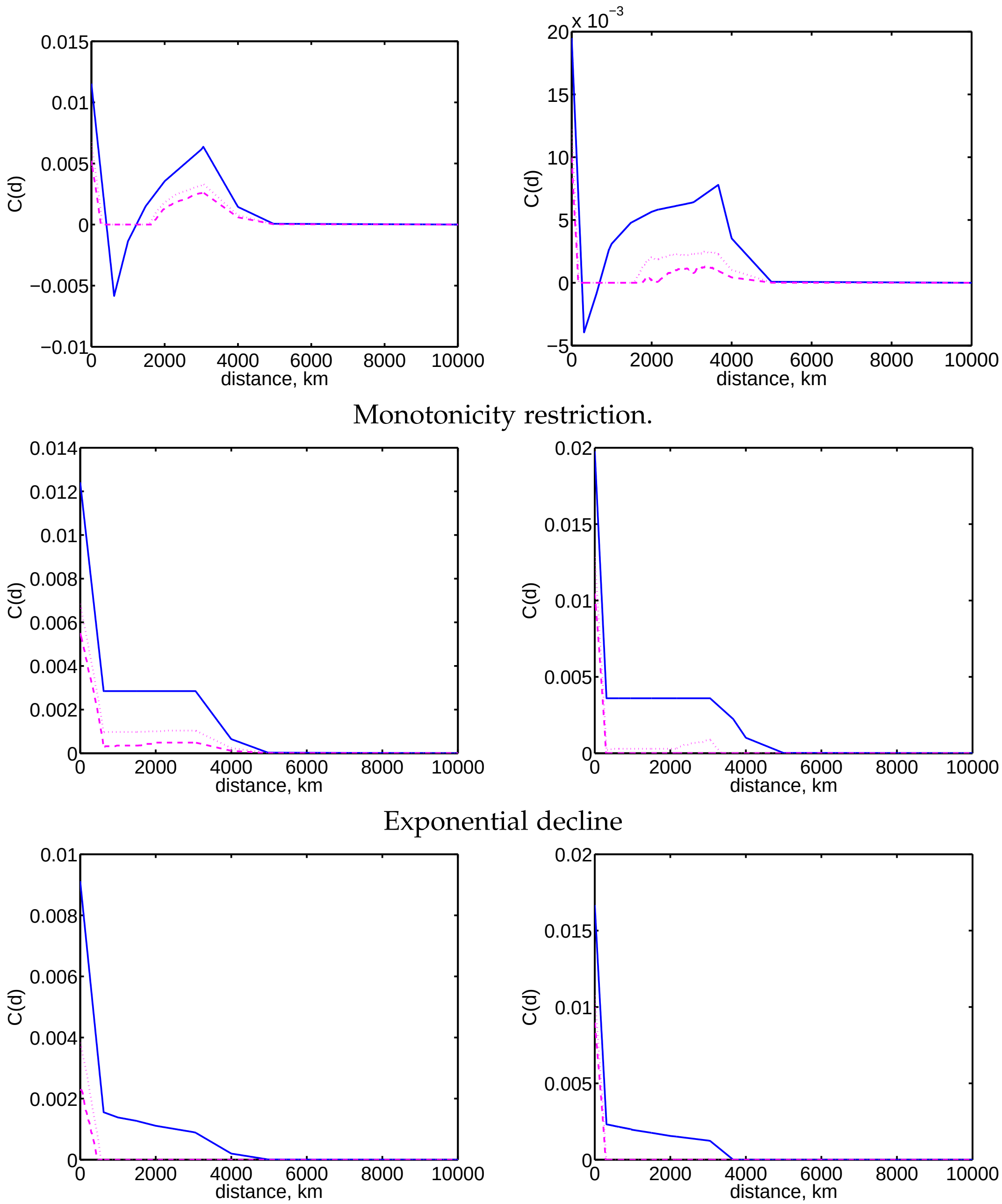
Figure 5 shows estimates of $C(d)$ from our semiparametric estimations under various specifications of the approximating function. The three rows of graphs correspond to the three sets of restrictions, or lack thereof: no restrictions, monotonicity, and exponential decline. The graphs on the left side were obtained by assuming that J_n = 2 (5 basis functions), whereas those on the right assume that J_n = 3 (9 basis functions). Table 9 lists the estimates of the parametric parts of the equation.

As is typical with series estimation, the unrestricted functions are highly nonmonotonic.

Figure 5: Estimates of the distance weight.

Five basis functions ($J_n = 2$) | Nine basis functions ($J_n = 3$)

No restrictions.



Notes:
The rate of decline is restricted to be at least 50% per 3000km. Measure of match: product. Dotted and dashed lines depict the boundaries of the 90% and 95% one-sided confidence bands.

Here, moreover, failing to impose exponentially declining functions is likely to result in inconsistent estimates; see Pinkse and Slade (2010). The other functions are, of course, monotonic by assumption. Regardless of specification, however, one regularity stands out — the function declines sharply close to the origin. Indeed, after about 500 km., geographic spillovers are small or nonexistent. This finding implies that technology markets are confined to narrow geographic areas.[14]

7.4 Correction for Endogeneity

IV Estimates

An obvious concern with the estimation strategy is that even after controlling for fixed effects the production function inputs are endogenous. There is a huge literature on estimating production functions (see, e.g. Ackerberg et al, 2007). We present some experiments in table 10 to allow for endogeneity following Arellano and Bond (1991). Column (1) presents estimates of the static production function coefficients. The results are consistent with our main within–groups estimates: geographically–based knowledge spillovers are positive and significant, albeit with a coefficient that is somewhat smaller than the Jaffe–based technological distance measure. As before, the product–market–based spillover term is statistically insignificant.

The diagnostics at the base of column (1) indicate that there is significant autocorrelation, which invalidates our use of instruments. If we follow Blundell and Bond (2000) and assume that there is an AR(1) error term in the static production function, this generates a model that includes a lagged dependent variable and lags of all the production function inputs and spillovers. We present the results of the unrestricted model in column (3), and note that the diagnostics at the base of the column indicate that there is no further evidence of serial correlation.

As outlined in section 6.2, the structure of the model implies common factor (COMFAC) restrictions on the coefficients, which we impose in column (2). The qualitative results from this COMFAC model are consistent with the findings in column (1) and elsewhere in the paper. In particular, there is a significant effect of the geographically–based R&D spillover in addition to the standard technological spillover term. We also experimented with using further "levels" moments that follow if we make assumptions over the initial conditions (essentially mean stationarity) as suggested by Arellano and Bover (1995) and Blundell and Bond (1998).[15] Unfortunately, Sargan–Difference tests suggested that these

[14]This pattern of local spillovers has also been found from patent citation data for universities (see Belenzon and Schankerman, 2010).

[15]Note, that unlike the main estimates, the coefficient on own R&D is no longer significant. This probably reflects the well-known problem discussed in Blundell and Bond (1998) that it is difficult to instrument very

Table 10: Productions functions Estimated by GMM

Dependent variable is ln(sales), $\ln Q_t$	(1)	(2)	(3)
	GMM	(restr.)	(unrestr.)
$\ln Q_{t-1}$			0.368
			(0.048)
lnK	0.149	0.192	0.192
	(0.052)	(0.044)	(0.048)
$\ln K_{t-1}$			-0.174
			(0.043)
lnL	0.619	0.546	0.514
	(0.058)	(0.051)	(0.054)
$\ln L_{t-1}$			-0.052
			(0.059)
lnOwnS	0.020	-0.021	-0.097
	(0.028)	(0.049)	(0.072)
$\ln OwnS_{t-1}$			0.066
			(0.047)
lnSpillSIC	0.012	-0.023	0.073
	(0.102)	(0.078)	(0.081)
$\ln SpillSIC_{t-1}$			0.008
			(0.063)
lnSpillTech	0.716	1.108	-1.464
	(0.303)	(0.313)	(0.903)
$\ln SpillTech_{t-1}$			2.159
			(0.887)
lnSpillGeog	0.690	0.609	0.401
	(0.276)	(0.262)	(0.564)
$\ln SpillGeog_{t-1}$			-0.075
			(0.573)
Autocorrelation		0.406	
		(0.046)	
LM(2) p–value	0.017		0.184
Hansen p–value	0.008		0.006
Comfac p–value		0.001	
Observations	9148	8260	8260

Notes:
Coefficients with standard errors in parentheses clustered by firm (these are the one-step robust results for GMM). All factor inputs and spillover terms are treated as endogenous and instrumented using the Arellano-Bond (1991) method using instruments dated t-2 to t-6 (i.e. all variables first differenced and instrumented with their own lagged levels from t-2 and before). Column (1) estimates a static production function and column (3) a dynamic production function including a lagged dependent variable and lags of all the variables. Column (2) imposes the Comfac restrictions i.e. it assumes that the DGP is a static production function with an AR(1) levels error term (see Blundell and Bond, 2000). All columns include a full set of time dummies.

moments are violated in our data implying that we cannot use the additional instruments.

persistent variables such as R&D stocks because lagged levels are poor predictors of future growth rates.

The estimates are far from ideal. First, Hansen/Sargan tests reject instrument validity in both specifications. This is common, however, when one has large samples (e.g. Nevo, 2000). Second, the COMFAC restrictions are rejected, which formally implies that we cannot go from the unrestricted to the restricted model. Nevertheless, we present our results and note that the positive and significant effect of the spillover terms is invariant to many experiments with different instrument sets. This finding suggests that our key conclusions are robust.[16]

Semiparametric IV Estimates

We have also estimated the COMFAC specification using our semiparametric estimator. The results are reported in table 11 and figure 6. Although there are differences with the unrestricted (i.e. without common factor restrictions) case, the main conclusions are the same.

Table 11: Semiparametric COMFAC estimates, parametric part

Dependent variable is ln(sales)						
	(1)	(2)	(3)	(4)	(5)	(6)
Restrictions	none	mono	exp decl	none	mono	exp decl
# of basis functions	5	5	5	9	9	9
$\ln Q_{t-1}$	0.365**	0.377**	0.383**	0.388**	0.403**	0.409**
lnK	0.102**	0.104**	0.104**	0.131**	0.137**	0.139**
lnL	0.627**	0.622**	0.622**	0.568**	0.579**	0.582**
lnOwnS	0.057*	0.060*	0.063*	0.077**	0.080**	0.083**
lnSpillSIC	0.042	0.048	0.051	0.028	0.007	0.002
lnSpillTech	0.933**	1.001**	1.021**	0.834**	0.870**	0.867**

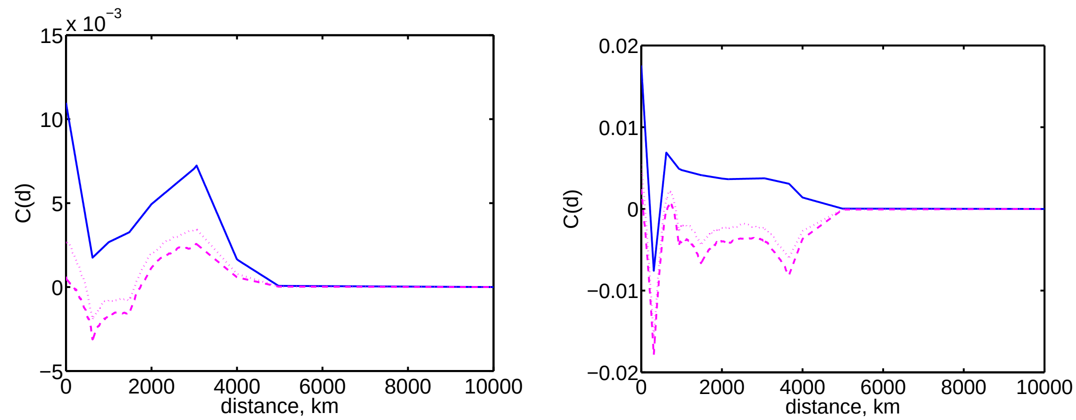
Notes:
One-step GMM is used
GMM weighting matrix is computed under the assumption of homoskedasticity
Moments in levels are not used
Differences are instrumented by the GMM-style lags in the range of 2 to 6
Exponential decline restriction requires that the rate of decay is at least 50% per 3,000 km
Year fixed effects included.
Inventor geographic measures, product measure of match.
Significance levels are obtained by subsampling
Number of subsamples: 1,000. Size of the subsample: 200 firms
Significance level: † – 10%, * – 5%, ** – 1%.

[16] We could go for an alternative dynamic form, say from a general to specific method. However, we do not want to overwhelm the reader with too many specifications. Indeed, there are diminishing returns to changing the specification.

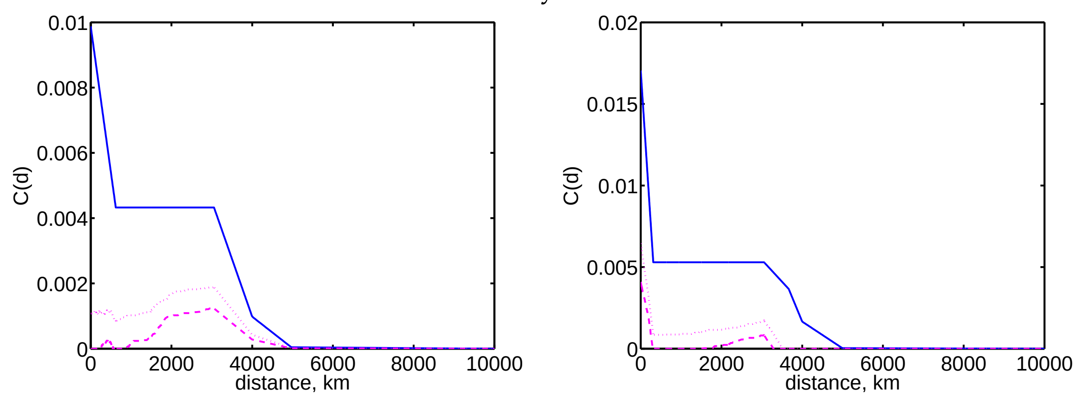
Figure 6: Estimates of the distance weight: Comfac.

Five basis functions	Nine basis functions
($J_n = 2$)	($J_n = 3$)

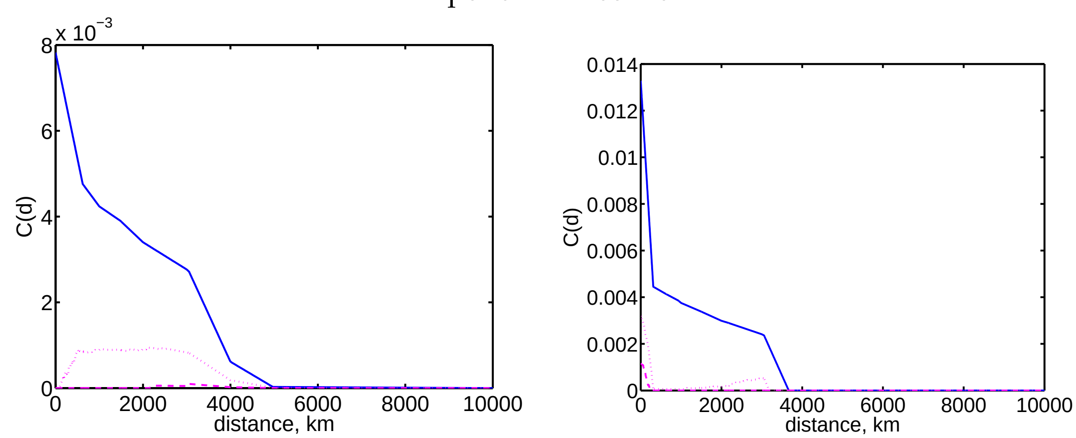
No restrictions.



Monotonicity restriction.



Exponential decline



Notes:
The rate of decline is restricted to be at least 50% per 3000km. Measure of match: product.
Dotted and dashed lines depict the boundaries of the 90% and 95% one-sided confidence bands.

8 Conclusions

A number of conclusions can be drawn from our study. First and foremost, geography matters. Indeed, intraregional spillovers are significant, sizeable, and economically important, even after conditioning on technological and product–market spillovers. This finding is consistent with an emerging literature suggesting important local productivity spillovers (e.g. Greenstone, Hornbeck and Moretti, 2010). We conclude that social learning and capitalization of complementarities among firms' research activities can be important factors for economic growth. Second, distance matters. More precisely, a geographic distance function that allows spillovers to decay gradually as regions become farther apart outperforms a specification that constrains spillovers to occur only within regions (i.e., a zero/one distance function that indicates that two regions are the same or different). Finally, estimated geographic spillover effects are larger when the distribution of each firm's inventors is used as a measure of R&D location rather than the location of its headquarters. Our use of the patent data to create spatial distributions of the location of firms' research activities therefore appears to be a worthwhile extension.

Turning to technological spillovers, we have experimented less with different measures of this important variable. Nevertheless, as with geographic spillovers, we find that technological spillovers are significant, sizeable, and economically important. In fact, in most specifications, the magnitude of the coefficients of this measure are comparable to those of the geographic measures.

The picture with respect to product–market productivity spillovers is very different. Indeed, although the coefficient of this variable is positive and significant when it is considered in isolation, more often than not that coefficient becomes negative and insignificant when the other spillover variables are included in the specification. We argue that this finding is consistent with the model outlined in Bloom et al (2008), where R&D by product market rivals affects market value, but should have no effect on total factor productivity, as it does not alter technological capabilities conditional on own R&D.

What have we learned from a policy point of view? We can conclude that since estimated spillover effects are large, there is a sizeable public–good aspect to R&D activity. In the absence of public policy to rectify this externality, we might therefore expect to see underinvestment in R&D. Nevertheless, as there are also costs associated with intervention, more research would be required before one could advocate interference in R&D markets. Finally, we can conclude that geographic markets are very local. This could explain why local policymakers invest substantial sums in tax incentives to attract R&D labs to their cities, states, and countries (e.g. Wilson, 2008). Indeed, our semiparametric estimates show that, although knowledge spillovers are large, they decay quickly, which is not good news for regional convergence.

Possible Extensions

Note that, although we have used specifications of proximity measures that allow for cross–market spillovers only for geographic markets, the same could be done for product and technology markets. To illustrate, consider product markets. Although it is clear that the production of cars is much closer to that of trucks and buses than it is to that of breakfast cereals and ladies apparel, the zero/one distance function that is implicitly assumed in research to date does not allow for intermediate cases (i.e., two activities are in different but close technology or product markets). It might therefore be fruitful to modify the product and technology proximity measures to incorporate nontrivial C functions. Moreover, the construction of SICs and technology classes, which involves classification into n–digit groups, gives us a natural distance metric.

References Cited

Ackerberg, Daniel, Caves, Kevin and Garth Frazer (2007) "Structural Identification of Production Functions," UCLA mimeo.

Acs, Z., Audretsch, D., and Feldman, M. (1992) "Real Effects of Academic Research: A Comment," *American Economic Review*, 82: 363–367.

Adams, J.D. (1990) "Fundamental Stocks of Knowledge and Productivity Growth," *Journal of Political Economy*, 98: 673–702.

Arellano, Manuel and Stephen Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* 58: 277–297.

Arellano, Manuel and Olympia Bover (1995) "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics* 68–1: 29–52.

Belenzon, Sharon and Mark Schankerman (2010) "Spreading the Word: Geography, Policy and University Knowledge," LSE/Duke mimeo

Bernstein, J.I. (1988) "Costs of Production, Intra– and Interindustry R&D Spillovers: Canadian Evidence," *Canadian Journal of Economics*, 21: 324–347.

Bernstein, J.I. and Nadiri, M.I. (1989) "Research and Development and Interindustry Spillovers: An Empirical Application of Dynamic Duality," *Review of Economic Studies*, 56: 249–269.

Bloom, N., Schankerman, M., and Van Reenen, J. (2008) "Identifying Technology Spillovers and Product Market Rivalry," CEP mimeo.

Blundell, R. and Bond, S. (2000) "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19: 321–340.

Chen, X. (2007) "Large Sample Sieve Estimation of Semi-nonparametric Models, chapter 76 in Handbook of Econometrics, Vol. 6B, James J. Heckman and Edward E. Leamer eds., North-Holland.

Dechevsky, L. and S. Penev (1997) "On Shape–Preserving Probabilistic Wavelet Approximators," *Stochastic Analysis and Applications,* 15, 187–215.

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

Eaton, J. and Kortum, S. (1999) "International Technology Diffusion: Theory and Measurement," *International Economic Review*, 40: 537–570.

Goto, A. and Suzuki, K. (1989) "R&D Capital, Rate of Return on R&D Investment, and Spillovers of R&D in Japanese Manufacturing Industries," *Review of Economics and Statistics*, 71: 555–564.

Greenstone, Michael, Rick Hornbeck and Enrico Moretti (2010) "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," forthcoming in the *Journal of Political Economy*.

Griffith, Rachel, Harrison, Rupert and John Van Reenen (2006) "How Special Is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing," *American Economic Review*, 96–5: 1859–1875.

Grilliches, Z. (1979) "Issues in Assessing the Contributions of Research and Development to Productivity Growth," *Bell Journal of Economics*, 10: 92–116.

Grilliches, Z. and Lichtenberg, F. (1984) "Interindustry Technology Flows and Productivity Growth: A Reexamination," *Review of Economics and Statistics*, 66: 324–329.

Grilliches, Z. (1991) "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, 94: 29–47.

Hall, Bronwyn, Jaffe, Adam, and Manuel Trajtenberg (2001) "The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools," NBER Working Paper 8498.

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

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

Jaffe, A.B. (1989) "Real Effects of Academic Research," *American Economic Review*, 79: 957–970.

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

Jaffe, Adam and Manuel Trajtenberg. (2002) *Citations and Innovations: A Window on the Knowledge Economy,* Cambridge: MIT Press

Keller, W. (2002) "Geographic Localization of International Technology Diffusion," *American Economic Review*, 92: 120–142.

Levin, R.C. and Reiss, P.C. 1988) "Cost–Reducing and Demand–Creating R&D with Spillovers," *RAND Journal of Economics*, 19: 538–556.

Levinsohn, J. and Petrin, A. (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70: 317–341. 60: 531–542.

Marschak, J. and Andrews, W.H. (1944) "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12: 143–205.

Olley, S. and Pakes, A. (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64: 1263–1298.

Orlando, M.T. (2004) "Measuring Spillovers from Industrial R&D: On the Importance of Geographic and Technological Proximity," *RAND Journal of Economics*, 35: 777–786.

Ornaghi, C. (2006) "Spillovers in Product and Process Innovation: Evidence from Manufacuring Firms," *International Journal of Industrial Organization*, 24: 349–380.

Pinkse, J., Slade, M.E., and Brett, C. (2002) "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70: 1111–1155.

Pinkse, J. and Slade, M.E. (2004) "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review*, 48: 617–643.

Pinkse, J. and M.E. Slade (2010) "The Future of Spatial Econometrics," *Journal of Regional Science*, 50: 103–117.

Politis, Dimitris N., Romano, Joseph P., and Michael Wolf (1999) *Subsampling*, New York: Springer.

Scherer, F.M. (1982) "Inter–Industry Technology Flows and Productivity Growth," *Review of Economics and Statistics*, 64: 627–634.

Spence, M. (1984) "Cost Reduction, Competition, and Industry Performance," *Econometrica*, 52: 101–121.

Stock, J. and M. Watson, (2008) "Heteroskedasticity-Robust Standard Errors for Fixed Effect Panel Data Regression," *Econometrica* 76, 155–174.

Villalonga, B. (2004) "Diversification Discount or Premium? New Evidence from the Business Information Tracking Series," *Journal of Finance* 59–2, 479–506.

Wilson, Daniel J. (2009) "Beggar thy Neighbor? The In–State, Out–of–State and Aggregate Effects of R&D Tax Credits," *Review of Economics and Statistics* 91: 431–436.

A Data Appendix

A.1 The patents and Compustat databases

The NBER patents database provides detailed patenting and citation information for around 2,500 firms (as described in Hall, Jaffe and Trajtenberg (2005) and Jaffe and Trajtenberg (2002)). We started by using the NBER's match of the Compustat accounting data to the USPTO data between 1970 to 2000[17], and kept only patenting firms. These firms were then matched into the Compustat Segment ("line of business") Dataset keeping only the 817 firms with data on both sales by four digit industry and patents, although these need not be concurrent. For example, a firm which patented in 1985, 1988 and 1989, had Segment data from 1993 to 1997, and accounting data from 1980 to 1997 would be kept in our dataset for the period 1980 to 1997. The Compustat Segment Database allocates firm sales into four digit industries each year using firm's descriptions of their sales by lines of business. See Villalonga (2004) for a more detailed description.

The output of each firm is deflated by the PPI for the primary industry of that firm. When PPI data are unavailable (about a third of the firm–year pairs in our sample) we use the aggregate Manufacturing PPI for those sectors.

This dataset was cleaned to remove accounting years with extremely large jumps in sales, employment or capital signalling merger and acquisition activity. When we removed a year we treat the firm as a new entity and give it a new identifier (and therefore a new fixed effect) even if the firm identifier (CUSIP reference) in Compustat remained the same. This is more general than including a full set of firm fixed effects as we are allowing the fixed effect to change over time. We also removed firms with less than four years of non-missing accounting data.

This left a final sample of 812 firms to estimate the model on with accounting data for at least some of the period 1980 to 2000 and patenting data for at least some of the period between 1970 and 1999. The panel is unbalanced as we keep new entrants and exitors in the sample.

A.2 Geographic coordinates

We obtain the geographic coordinates of inventors by merging the patents data and the extract from the Geographic Names Information System (GNIS) provided by the US Geological Survey. All inventor records contain address information, in most cases the names of the home state and the city. The GNIS records provide coordinates for all established geographic names.

[17] We dropped pre-1970 data as being too outdated for our 1980s and 1990s accounts data.

Most of the time city and state names are not sufficient to uniquely identify a geographic place: a typical city name has a number of duplicates within its state. If these duplicates are located in the same county or within 0.1 degrees of latitude and longitude, we treat them as one entity.

If the duplicated names are located far apart, we give preference to incorporated places. Names of incorporated places are usually more permanent and stable, they are more likely to be used by the U.S. Postal Service. In turn, inventors are more likely to indicate a USPS-preferred city name when they are asked to provide their home address.

Following these rules, we are able to match 98.25% of inventor records that are related to the firms in our sample. To reduce computational burden, we aggregate all geographic data to the county level. For example, we assign identical coordinates to inventors from Long Beach CA and Los Angeles CA: both cities are located in Los Angeles County.

A.3 Variables

The book value of capital is the net stock of property, plant and equipment (Compustat Mnemonic PPENT); Employment is the number of employees (EMP). R&D expenditure (XRD) is used to create R&D capital stocks calculated using a perpetual inventory method with a 15% depreciation rate. We use sales as our output measure (SALE). Industry price deflators were taken from the U.S. Bureau of Labor Statistics PPI Database.

The construction of the spillover variables is described in Section (3) above in detail. About 95% of the variance of the spillover measures, SpillSIC, SpillTech and SpillGeog is between firm and 5% is within firm. When we include fixed effects we are, of course, relying on the time series variation for identification. lnSpillGeog is normalized to have mean zero and variance one to facilitate comparisons of coefficients. In particular, when we compare the explanatory power of different measures of match and distance, we use the fact that when two variables have the same variance, the variation that they contribute is measured by the absolute values of their coefficients.

Industry sales are constructed in exactly the same way as SpillSIC, with the only exception that the market proximity measure w^P is weighted with the firm's sales instead of the R&D stock.