# ON THE IMPORTANCE OF GEOGRAPHIC AND TECHNOLOGICAL PROXIMITY FOR R&D SPILLOVERS: AN EMPIRICAL INVESTIGATION


**Michael J. Orlando**
Federal Reserve Bank of Kansas City
Economic Research Department
925 Grand Boulevard
Kansas City MO 64198
phone: 816-881-2695
fax: 816-881-2199
e-mail: michael.j.orlando@kc.frb.org


October 2000


Abstract: Empirical studies of the external effects of R&D suggest that both geographic and technological distance attenuate inter-firm spillovers. The results presented here indicate that degree to which R&D spillovers localize economic activity is dependent on the technological relation between spillover generating and receiving firms. A production function framework is employed which controls for correlation between measures of geographic and technological proximity. Coefficient estimates confirm that R&D spillovers are largest among technological neighbors. Spillovers within narrowly defined technological groups, however, do not appear to be attenuated by distance. Geographic proximity serves to attenuate only those inter-firm spillovers that cross narrowly defined technological boundaries.



*Key words*: R&D, spillovers, industrial agglomeration, geography, empirical studies.
*JEL classification*: O3, L6



The author is an economist at the Federal Reserve Bank of Kansas City. The views expressed in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System. Omissions or errors are the sole responsibility of the author and may persist in spite of the invaluable guidance and comments provided by Bruce Petersen, Marcus Berliant, Steven Fazzari, Douglass North, Robert Parks, Josh Lerner, and the participants of the NBER Summer Institute 2000: Productivity, Science, and R&D.

## 1. INTRODUCTION

Location in geographic and technological space are two factors that may affect the ability of a firm to internalize the results of another firm's research and development (R&D) activity. Inter-firm R&D spillovers are presumed to be larger if firms are in close geographic proximity. Firms that are technologically similar are also presumed to have greater occasion to exchange spillovers. Identification of the importance of each of these factors is compromised, however, if firms in the same industry agglomerate for reasons exclusive of localized inter-firm knowledge spillovers.

The purpose of this paper is to examine the importance of geographic and technological factors for R&D spillovers. The study also considers the impact of industrial agglomeration on evidence relating to the importance of these factors. Direct evidence of spillovers is obtained from a firm-level production function framework in which geographically and technologically proximate R&D stocks are found to be correlated with output.

Results from a preliminary analysis are consistent with the stylized facts that geographic and technological proximity enhance R&D spillovers. The final analysis, which controls for the extent of industrial agglomeration in these data, indicates that the magnitude of R&D spillovers and the degree to which they are localized depends upon the technological relation between spillover producing and receiving firms. Spillovers are largest among firms *within* the same narrowly defined industry but are insensitive to inter-firm distance. To the extent that firms engaged in technologically similar industrial activity agglomerate geographically, they appear to do so for reasons exclusive of localizing intra-industry knowledge spillovers.

R&D spillovers *across* narrowly defined industry boundaries differ from their intra-industry counterparts in two respects. They are smaller in magnitude than those from technologically similar firms. However, these spillovers appear to be sensitive to geographic proximity, suggesting knowledge spillovers may serve a localization function for a diversity of industrial activity. Failure to control for the propensity of technologically similar firms to cluster in space may lead researchers to overstate the importance of geographic proximity for spillovers, particularly for those among technologically similar firms.

The next section places R&D spillovers into context by summarizing the range of areas of economic inquiry in which R&D spillovers play a prominent role. Several widely cited contributions to the measurement of inter-firm spillovers are surveyed. In section three I extend the production function framework recommended by Hall and Mairesse (1995) to include external stocks of R&D. In section four, county-specific latitude/longitude data is merged with a panel of annual financial variables to characterize the relative proximity of each pair of firms in both geographic and technological dimensions. Empirical results are summarized in sections five and six. The implications of these findings are discussed in the concluding section.

## 2. MOTIVATION AND LITERATURE REVIEW

Imperfect appropriability conditions distinguish R&D activities from normal factors of production. Whereas service flows from conventional inputs are fully appropriable, service flows from R&D expenditures have public good-like qualities. Ideas for new products and production processes are non-rival; firms regularly make multiple and simultaneous use of inventions from their own R&D laboratories. And while patents and trade secrets are routinely used to exclude others from use of one's ideas, enforcement is costly and of limited effect. This positive externality to private investments in knowledge creation plays an important role on a range of frontiers of economic inquiry, including theories of firm performance and industrial structure, regional development, and economic growth.

Schumpeter (1942) proposed that incentives to invest in innovation will be maximized by industrial structures that shelter potential innovators from the "perennial gale of creative destruction." Spence (1984) formalized the tradeoff whereby static inefficiency associated with market power is offset by improved dynamic efficiency associated with increased appropriation of returns to innovation. Thus, fundamentals that govern inter-firm spillovers ultimately determine appropriability conditions that govern industrial structure and performance.

City formation is hypothesized to arise in part from externalities, including those related to knowledge creation. This mechanism results in localization, or the agglomeration of similar industrial activity, if knowledge spillovers are confined to firms in the same industry. Urbanization, or the agglomeration of a diversity of industrial activity, will occur if knowledge spills across industry boundaries.[1] In either case, agglomeration will only occur if distance attenuates knowledge spillovers.

Finally, knowledge intensive activities such as R&D play a central role in models of economic growth constructed in both the Arrow 'learning-by-doing' and Schumpeterian traditions.[2] Following from the dual private/public nature of knowledge, the output of research activity is understood to be non-rival and (at least partly) non-excludable. Thus, private investments in knowledge creation may augment others' productivity, allowing the economy to sufficiently avoid diminishing returns at the aggregate level for some sets of parameter values. Understanding the nature of the externality to knowledge creation is therefore central to understanding whether such models provide useful insights for explaining observed economic growth.

A growing body of empirical evidence suggests spillovers are significant and typify the appropriability conditions of R&D. The earliest analyses find social returns exceed private returns to investments in innovative activity.[3] Responses to a survey of R&D managers indicate that spillovers are an important source of ideas for innovation intensive firms.[4] And surveys of regression based estimates suggest R&D spillovers are important at both the inter-industry and inter-firm levels.[5]

---

[1] Localization externalities are sometimes referred to as MAR externalities after Arrow's (1962) and Romer's (1986) formalization of Marshall's (1890) view that the economically most important spillovers are among firms in the same industry. Industrial concentration facilitates regional economic growth by enhancing the realized benefits of interfirms spillovers. This view is often contrasted with urbanization externalities, sometimes referred to as Jacobs externalities, after the view advanced by Jacobs (1969) that economically significant spillovers occur through application of knowledge beyond the industry of origin (Glaeser et al. 1992 p. 1127). In this case, industrial diversity facilitates regional economic growth.

[2] See Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992).

[3] See Griliches (1958), Terleckyj (1958), and Minasian (1962).

[4] Levin et al. (1987) list a variety of ways a firm may learn of rivals' technologies including patent disclosures, technical publications and conferences, high-tech labor market turnover, and reverse engineering of others' products.

[5] Surveys are provided by Cameron (1996), Nadiri (1993), and Griliches (1992).

Industry-level studies of U.S. data suggest returns to R&D appropriated outside the industry of origin are at least as large as those appropriated by firms from within. Moreover, own-industry returns to R&D can reflect a combination of direct returns and inter-firm spillovers within an industry. Indeed, in firm-level analyses, Jaffe (1986) and Bernstein and Nadiri (1989) find inter-firm spillovers from R&D are significant and important. Increasingly, attention has been focused beyond simply measuring R&D spillovers to an examination of the factors which facilitate their transmission.

Firms that are neighbors may exchange knowledge through a variety of channels. Knowledge may be transmitted through employee interaction in social, civic, and professional organizations, participation in which may be geographically constrained.[6] And routine employee turnover can result in significant cross-pollination of knowledge stocks. Geographically near firms are also likely to share buyers and suppliers who may serve as additional conduits for information flow.

Jaffe et al. (1993) provide evidence of localization of knowledge spillovers from inventive activity. The authors find that a patent's citations are more likely to come from the same state and standard metropolitan statistical area than are a technologically similar control group of patents. Audretsch and Feldman (1996) suggest this localization of knowledge spillovers applies to innovative activity as well. They find greater clustering of innovative activity in industries where knowledge externalities are likely to play a greater role. They interpret concentration of innovation in excess of geographic concentration of production as evidence in support of the hypothesis that geographic distance attenuates knowledge spillovers.

Empirical evidence from the urban/regional literature is also consistent with the presence of geographically dependent knowledge spillovers. Glaeser et al. (1992) find that local competition and urban variety encourage employment growth in cities, a finding consistent with the view advanced by Jacobs (1969). Henderson et al. (1995) find that the industrial concentration/employment growth relation is dependent on industry life-cycle effects. They argue that firms in young industries thrive in industrially diverse cities

[6] See Saxenian (1994).

while firms in older, established industries benefit from localization, a conclusion consistent with the Marshall-Arrow-Romer (MAR) hypothesis.[7]

Spillovers are also believed to be higher among "technological neighbors". According to this view, the ability to make productive use of another firm's knowledge depends on the degree of technological similarity between firms. Every technology has a somewhat unique set of applications and language. Researchers in similar technological fields will interact in professional organizations, publish in commonly read journals, and, increasingly, browse a common set of web pages. Reverse engineering may be employed to maintain parity with one's rivals. And spying and corporate espionage are thought to be relatively common among information intensive industries.[8]

Jaffe (1986) uses firm panel data and patent histories to construct a technological proximity-weighted spillover pool for each firm.[9] The author finds evidence of technologically dependent spillovers in that firms active in research intensive technology groups enjoy higher research productivity and higher returns to R&D, controlling for industry specific technological opportunities.

Adams and Jaffe (1996) estimate both the geographic and technological sensitivity of spillovers at the intra-firm level. Using plant-level data, the authors find evidence of attenuating effects associated with both measures of distance. Specifically, they find that a firm's research lab activity has the greatest impact on its own plants within 100 miles and those working in a similar product class. Due to data limitations, however, the authors do not consider the two hypotheses simultaneously and are limited to a series of partial analyses.[10] Furthermore, while we expect firms to act to maximize spillovers among their own plants, it is likely that they try to minimize inter-firm

[7] See footnote 1.
[8] For an example, see "Oil Firms Still Rely on Corporate Spies to be Well Informed", in *The Wall Street Journal*, December 7, 1998.
[9] The author defines a multi-dimensional technology space spanning 47 individual patent classes. A vector in this space summarizes a firm's patent history and thereby defines their technological position. The weight of one firm's R&D in the another's spillover pool is inversely proportional to the angular separation between patent vectors.
[10] Adams and Jaffe (1996, p. 703) note that partial analyses which find, for example, that technological distance mitigates spillovers may be detecting the importance of geographic proximity.

spillovers. Thus, it is not clear that the authors' intra-firm level findings can be generalized to spillovers across firm boundaries.

## 3. DEVELOPMENT OF HYPOTHESES

As a point of departure I follow Hall and Mairesse (1995) and assume firm output, $Y_{it}$, can be represented with a conventional Cobb-Douglas production technology as follows:[11]

$$Y_{it} = A\hat{K}_{it}^{\pi} e^{\lambda t} C_{it}^{\alpha} L_{it}^{\beta} e^{\varepsilon_{it}} \tag{1}$$

where $C_{it}$ is capital, $L_{it}$ is labor, $\alpha$ and $\beta$ are their respective output elasticities, and and $\varepsilon$ is an i.i.d. disturbance. The function $A\hat{K}_{it}^{\pi} e^{\lambda t}$ represents a firm-specific 'state of technology' which summarizes all knowledge relevant to firm *i*'s production possibilities in time period *t*.

The state of technology is a function of an exogenous time trend, $\lambda$, and other exogenous factors, $A$.[12] Assuming R&D is the single determinable component of knowledge relevant to the production process, then $\hat{K}_{it}$ represents the effective stock of R&D contributing to firm *i*'s production in time period *t*. If the outcome of R&D efforts are at least partly non-rival and non-excludable then $\hat{K}_{it}$ will be an aggregation of R&D stocks accumulated through one's own research effort, as in Hall and Mairesse (1995), and those available through spillovers from other firms.

Construction of this effective stock of R&D has been the focus of many early

[11] While one may argue that little can be learned about spillovers from the partial equilibrium relationship summarized by a firm production function, Hall and Mairesse (1995) find that the production function framework is preferred to the rate of return formulation. Furthermore, through comparison to the 'semi-reduced form' approach, the authors find the production function does not yield biased estimates of R&D elasticity when controls for permanent firm effects are included. One may argue still for something more behavioral, such as a location theoretic framework. But why firms are where they are is presumably the outcome of utility maximizing decisions of decisionmakers. The purpose of this paper is to say something about the way in which spillovers occur and so it seems reasonable, before complicating the analysis with complex behavioral stories, to ask just what insights a simple model of spillovers might yield.

[12] For ease of exposition, the assumption of a common intercept term, $A$, and time trend, $\lambda$, is maintained throughout the derivation. The estimation procedure relaxes this assumption by incorporating state, industry, and year specific intercept terms. The assumption of constancy in the other parameters should not be too offensive since the scope of the analysis will be limited to a single two-digit standard industrial classification.

attempts to identify the presence of inter-firm spillovers. As we turn our attention to understanding the nature of R&D spillovers, such measurement issues may be recast as empirical questions. In contrast to earlier studies which have assumed a weighting scheme consistent with spillover channels described in section 2, the strategy pursued here is to allow the data to reveal the structure of this weighting scheme.

Assume that the effective stock of R&D for firm *i* can be expressed as follows:

$$\hat{K}_{it} = rK_{it}^{-\rho} + \sum_{j \neq i}^{I} s_{ij} K_{jt}^{-\rho} \tag{2}$$

where $K_{it}$ and $K_{jt}$ represent R&D stocks employed by the own firm and parties external to the firm, respectively. Substitution of (2) into (1) and redefining $\pi = -\frac{1}{\rho}$ leaves a production function in which individual R&D stocks are complimentary to conventional inputs but interact with one another in a familiar constant elasticity of substitution (CES) process:

$$Y_i = Ae^{\lambda t} C_i^{\alpha} L_i^{\beta} (rK_i^{-\rho} + \sum_{j \neq i}^{I} s_{ij} K_j^{-\rho})^{-\frac{1}{\rho}} e^{\varepsilon_i} \tag{3}$$

where the subscript *t* is suppressed for clarity. In principle, then, each firm, *i*, may have access to a unique state of technology. This technology is a combination of own R&D with weight *r* and the R&D stocks of other firms, *j*, each with weight $s_{ij}$.

It is well known that a special case of the CES production function is that where $\rho$ = 0, corresponding to a Cobb-Douglas technology. A second order Taylor series expansion of (3) around that point may be written as

$$y_i = a + \lambda t + \alpha c_i + \beta l_i + rk_i + \sum_{j \neq i}^{I} s_{ij} k_j - \frac{\rho}{2} \sum_{j \neq i}^{I} rs_{ij} (k_i - k_j)^2 - \frac{\rho}{2} \sum_{h \neq i}^{I} \sum_{j \neq i}^{I} s_{ih} s_{ij} (k_h - k_j)^2 + \varepsilon_i \tag{4}$$

where lower-case variables represent natural logarithms of their upper-case counterparts.[13] The $s_{ij}$'s are interpreted as elasticities of output with respect to the R&D

[13] The linear case ($\rho$ = -1) was also estimated, however, goodness of fit was reduced as compared to the Cobb-Douglas case and spillover pool coefficients were not significant. Admittedly, this could be a consequence of the noisier nature of the spillover pool variable in this case. However, it may also be interpreted as evidence against the linear functional form which assumes own and others' R&D serve as

stock of the respective firm *j*. In principle, this linear regression analysis can be used to estimate the elasticities of output with respect to the R&D stocks of other individual firms. We could then examine the pattern of weights in this matrix and ask if they are reflective of the pattern of geographic and technological distance between firms. Unfortunately, direct estimation of this spillover weighting matrix is far too costly in terms of degrees of freedom. We can, however, appeal to *a priori* expectations regarding the pattern of variation in the $s_{ij}$'s to structure an estimable equation from the cumbersome reduced form of this most general case.

Consider the hypothesis that spillovers are a function of geographic distance between firms. For the purpose of this discussion, assume that the productivity of external R&D varies only with distance between the spillover sending and receiving firms. Specifically, the R&D stocks of nearby firms yield greater spillovers than do those of distant firms.[14] In order to make (4) empirically manageable, we need to limit the number of parameters requiring estimation. The technique used here is to draw a circle of arbitrary radius, *R*, around each firm *i* and define all firms inside that circle elements of the set $G_{Ni}(R)$. All firms outside of the circle are considered elements of the set $G_{Zi}(R)$. Equation (4) may then be expressed as

$$y_i = a + \lambda t + \alpha c_i + \beta l_i + r k_i + s_N \sum_{n \neq i}^{G_{Ni}} k_n + s_Z \sum_{z \neq i}^{G_{Zi}} k_z + \varepsilon_i \tag{5}$$

where the second order terms are dropped since $\rho$ is assumed to be 0. Our principal interest is with the magnitude and sign of the near spillover pool coefficient relative to that of the distant spillover pool and relative to the definition of spillover pool boundaries. If spillovers are sensitive to the relative locations of other innovating firms then the near R&D pool variable will be positively correlated with firm value added, $s_N >$

---

substitutes. This interpretation is consistent with Cohen and Levinthal (1989) who argue that incentives to engage in own R&D include development of absorptive capacity, or the ability to make use of technology developed by others. Furthermore, preliminary grid-search results from non-linear estimation of (3) suggest $\rho$ converges to a value between -0.1 and 0.1.

[14] Alternatively, one could argue that spillovers from nearby firms are received *faster* than those from distant firms. But if this is the case then lead time advantages (Levin et al. 1987) will make later service flows from external R&D less valuable the spillover receiving firm and the empirical effect should be the same.

0. The near R&D pool coefficient will also be larger than that of the distant spillover pool, i.e. $s_N > s_Z$. Furthermore, increasing the distance to the near pool / distant pool boundary will include firms farther and farther from firm $i$ into the set $G_{Ni}$, imposing a downward bias on $s_N$, i.e. $\frac{\partial s_N}{\partial R} < 0$.

Of course, this same apparatus may be used to evaluate the sensitivity of interfirm spillovers to technological similarity. Firms engaging in different research, using dissimilar technical processes, or making distinct goods may be considered "technologically distant" from one another. Diffusion of R&D services from one such firm to the next should be lower relative to diffusion between firms that are "technologically near". Therefore, tests of a technological proximity hypothesis analogous to those proposed in the preceding paragraph may be constructed by sorting all firms into technologically near and technologically distant pools.

As noted in section 1, if technological and geographic proximity are highly correlated then interpretation of results from the partial analyses described in the preceding paragraphs may be problematic. Therefore, the main analysis generalizes the pooling technique described above to simultaneously control for geographic and technological proximity. R&D stocks external to each firm $i$ are sorted into pools of firms that are geographically near/technologically near, geographically near/technologically distant, geographically distant/technologically near, and geographically distant/technologically distant. The resulting reduced form analogous to (5) is:

$$y_i = a + \lambda t + \alpha c_i + \beta l_i + rk_i + s_{NN} \sum_{n \neq i}^{G_N T_{Ni}} k_n + s_{NZ} \sum_{u \neq i}^{G_N T_{Zi}} k_u + s_{ZN} \sum_{v \neq i}^{G_Z T_{Ni}} k_v + s_{ZZ} \sum_{z \neq i}^{G_Z T_{Zi}} k_z + \varepsilon_i \qquad (6)$$

where $G_N T_{Ni}$, $G_N T_{Zi}$, $G_Z T_{Ni}$, and $G_Z T_{Zi}$ are the respective subsets of spillover generating firms for each firm $i$ and $n$, $u$, $v$, and $z$ are their respective indexes. It should therefore be noted that estimating (5) under the geographic (technological) proximity hypothesis is equivalent to estimating the more general model (6) with the restriction that $s_{NN} = s_{NZ}$ and $s_{ZN} = s_{ZZ}$ ($s_{NN} = s_{ZN}$ and $s_{NZ} = s_{ZZ}$).

The preceding derivation is based on two key assumptions. First, as stated above, the elasticities of own output with respect to external R&D, the $s_{ij}$'s, are assumed to vary only with distance between spillover producing and receiving firms. Returning to (4), this implies that $s_{ij}$ varies in $j$ only with the distance between firms $i$ and $j$. Second, implicit in the derivation of (5) is the assumption that the $s_{ij}$'s are also constant across spillover receiving firms, i.e. $s_{ij}$ is constant for all $i$ for a given pool boundary radius $R$.

It should be noted that the reduced form presented in (5) yields a spillover pool elasticity of output that is not scale neutral in the way of conventional input elasticities. The scale neutrality of conventional input elasticities stems from the assumption that the scale of inputs is correlated with the scale of outputs for a given firm size. For example, an elasticity of output with respect to capital of 0.1 is interpreted to mean that a one-percent increase in capital will yield a 0.1 percent increase in output, *ceteris paribus*. For a small firm, a one percent increase in capital will imply a 0.1 percent increase in output which is presumably small in absolute terms. Larger firms will experience a larger nominal change in output in proportion to their nominally larger one percent change in capital.

But firms share the same nominal size of spillover pool regardless of their own size. While the elasticity, $s$, is still technically defined as the percentage change in output attributable to a percentage change in spillover pool R&D, the scale of this R&D is constant across firms of varying size. This construction implicitly assumes that large firms receive a relatively larger benefit than small firms from the same size spillover pool.

Given that it is the non-rival nature of R&D that motivates this research in the first place, this assumption is intuitively appealing. Additional support is drawn from Jaffe (1986) who reports benefits from R&D spillovers increasing in own scale of R&D. Additional evidence in support of this implicit assumption is found in Henderson and Cockburn (1996) who conclude in an examination of size and research productivity in the pharmaceutical industry:

> "[L]arger research efforts are more productive … because they realize economies of scope by sustaining diverse

> portfolios of research projects that capture internal and external knowledge spillovers."[15]

Thus, the assumption that large firms benefit more from the same nominal base of spillover pool R&D seems sensible.

## 4. DATA SELECTION AND EMPIRICAL SPECIFICATION

This examination of inter-firm R&D spillovers is intended to illuminate the microfoundations presumed to drive models of urban/regional agglomeration and economic growth at the macro level. For several reasons then, SIC 35, Industrial, Commercial Machinery, and Computer Equipment, represents a prime candidate for analysis.[16] Innovations in these investment goods manufacturing industries are thought to be prime sources of economy wide growth in both the Arrow learning-by-doing and Schumpeterian traditions. Furthermore, selection of SIC 35 incorporates 357, Computer and Office Equipment, a sector increasingly perceived to play a central role in the renewed trend of productivity growth.[17]

SIC 35 is composed of the investment goods manufacturing industries listed in

**Table 1**
**SIC 35: three-digit standard industrial classifications**

| SIC | Name |
|---|---|
| 352 | Farm & garden machinery and equipment |
| 353 | Construction, mining, and material handling machinery and equipment |
| 354 | Metalworking machinery and equipment |
| 355 | Special industry machinery, except metalworking machinery |
| 356 | General industrial machinery and equipment |
| 357 | Computer & office equipment |
| 358 | Refrigeration & service industry machinery |
| 359 | Miscellaneous industrial and commercial machinery and equipment |

[15] Henderson and Cockburn (1996), p. 32.
[16] A review of the National Science Foundation publication *Research and Development in Industry: 1996* indicates that SIC 35 is one of the four most R&D intensive two-digit classifications; the others being 36 (Electrical Equipment), 37 (Transportation Equipment), and 28 (Chemicals and Pharmaceuticals). These four R&D-intensive industries account for nearly 60% of total privately funded R&D.
[17] See Jorgenson and Stiroh (2000) and Oliner and Sichel (2000).

table 1. Annual firm-level financial data used in the analysis are obtained from Standard & Poor's COMPUSTAT database. Input and output deflators are obtained from the Bureau of Labor Statistics website. County-level latitude and longitude data are obtained from the U.S. Geological Survey Geographic Names information System.

Methods of variable construction are summarized in table 4.2. Firm-level material and energy expenditures were not available. Estimation error imposed by use of sales as a proxy for output, however, will be confined to the constant term if these charges are a fixed proportion of sales.[18] This assumption is implicit in the firm fixed-effects model. Industry- and state-specific dummies will serve as controls in the cross-sectional regression to the extent that variation in the materials and energy fraction of sales is an industry- or region-fixed effect.

**Table 2**
**Summary of Variable Construction**

| Variable Symbol | Variable Name | Proxy |
|---|---|---|
| *Y* | Output | Sales |
| *C* | Capital | Estimated by accumulating capital spending following Salinger and Summers (1983) with correction for acquisitions and divestitures per Chirinko et al. (1999) |
| *L* | Labor | Number of employees |
| *K* | R&D stock | Estimated by accumulating R&D expenditure using perpetual inventory method assuming 15% depreciation rate and pre-sample period growth rate of 5% |

In the strictest sense, sales revenue will only be a sensible proxy for output where perfect competition obtains. In this case, real prices are assumed to be constant so all changes in total revenue are attributed to changes in real output. This assumption becomes problematic where firms enjoy market power. In this case, changes in total revenue may be attributed to actual output changes or changes in the real price a firm charges its customers. But if market power is assumed to result in a constant mark-up

[18] Basu (1996) reports that material inputs are nearly perfectly correlated with output.

then estimation error will be confined to the constant term.  While it is certainly arguable that this mark-up is constant across all observations in the panel it may be so across firms in a particular industry.  In this case, variations in market power will be controlled for with industry dummies.  Assuming market power is at least constant across time for a given firm, estimation error associated with market power will be controlled for though use of the firm fixed effects regression model.

The capital variable makes use of widely available accounting data on capital expenditures and is an improvement on the commonly used Salinger and Summers (1983) algorithm.[19]  Following Chirinko et al. (1999), I begin with the reported value for net property, plant, and equipment and iteratively accumulate and depreciate annual capital expenditures.  Firm year-end capital stocks are adjusted to account for divestitures and acquisitions.  Use of this algorithm avoids exclusion of firms engaged in mergers over the sample period, reducing sample size and introducing selection bias.

The R&D stock variable for each firm is constructed from R&D expense[20] using the perpetual inventory method commonly employed in studies of R&D productivity.[21] The results reported below are invariant to a preliminary check of sensitivity to various depreciation and pre-sample period growth rates.[22]

The labor variable is reported in thousands of employees.  All other financial variables are reported in millions of nominal dollars.  Nominal values are deflated with price indexes obtained from the Bureau of Labor Statistics website.  Sales figures are deflated with three-digit SIC-specific producer price indexes.  Capital is deflated with a capital expense index.  R&D stock is deflated with the occupational cost index for

---

[19] Chirinko et al. (1999).
[20] Cockburn and Griliches (1988) conclude that "… [d]ata on R&D expenditures … are stronger measures of input to the process by which firms produce technical innovation than patents are of its output."
[21] See Griliches (1979).
[22] Preliminary estimates of the results reported in the following sections were insensitive to depreciation rates ranging from five to 30 percent.  Hall & Mairesse (1995) also observe "… the choice of depreciation rate in constructing R&D capital does not make much difference to the coefficient estimates …  This result has already been observed in a number of previous studies and arises from the basic fact  that the time series of R&D expenditure *within firm* does not vary all that much."

technical professionals in order to reflect the fact that the majority of R&D expenditures represent staff salaries.[23]

Adding a geographic component to the analysis requires assigning each firm to a particular point in space. Firms are assumed to be located at the geographic centroid of the county location of their corporate headquarters. An arbitrary radius is chosen and the implied circle around each firm defines all other firms in the panel as geographically near (inside the circle) or geographically distant (outside the circle).[24]

Use of corporate headquarters to represent firm location may be questionable for the purpose of spillover detection. One may argue that our true interest is in the location of innovation, not necessarily in the location of corporate headquarters. However, if firms view R&D as their most strategically important investment they are likely to locate this activity close to corporate headquarters, suggesting little error will be introduced with the location assumption proposed here. Furthermore, while R&D may be a reasonable proxy of the scale of a firm's innovative activity, spillovers from this implied knowledge base may emerge from any of the locations that compose the firm; R&D facilities, production facilities, or corporate headquarters. Thus, corporate headquarters may be as good a proxy of firm location as we can hope to find.

Nevertheless, the convention in this literature is to define the location of R&D activity as the probable location of spillover generation and reception. Therefore, the

---

[23] According to the "Advanced Release of Selected Tables from the Research and Development in Industry: 1995" report by the NSF, a majority of R&D expense is for salaried technical professionals with materials and supplies accounting for 10% to 20%. Also, see Grabowski (1968). In the absence of an R&D specific price deflator I use the occupational cost index for technical professional.

[24] Each firm's geographic location is defined with the state and county name as reported in the COMPUSTAT Name and Location datafile. The COMPUSTAT Names and Financial Variables datafiles are merged with data obtained from the U.S. Geological Survey (USGS) Geographic Names Information System (GNIS). Each observation in this dataset reports the latitude and longitude of the geographic centroid of a county in degrees, minutes, and seconds. Federal Information Processing Standard state and county codes are used to map a set of latitude/longitude coordinates from the USGS GNIS database to each observation in the COMPUSTAT Financial Variables datafile. The distance between any two firms in a given year is then computed as the distance between their respective county centroids. Assuming a spherical earth of actual earth volume, the arc distance in miles between any two points *i* and *j* can be derived as $d_{ij} = 2 * 3{,}959 * \arcsin \sqrt{\sin^2\left(\frac{lat_j - lat_i}{2}\right) + \left(\frac{\cos(lat_j) + \cos(lat_i)}{2} \sin\left(\frac{lon_j - lon_i}{2}\right)\right)^2}$ where 3,959 is the radius of the earth in miles and latitude and longitude values are in radians.

*Directory of American Research and Technology 1993* was consulted to establish the reasonableness of the claim that corporate headquarters may be a useful proxy for the location of R&D activity. This volume catalogs the location of firm's corporate headquarters and the location and composition of activity at each of that firm's R&D facilities. Approximately half of the firms included in this analysis have entries in the directory. Of this sample, 87 percent of firms conduct at least part of their R&D in the same city as their corporate headquarters. And fully 73 percent of the sample of firms conduct their R&D exclusively in the same city as their corporate headquarters. These sample statistics are likely to represent a lower bound if we assume that any underreporting in this directory is biased towards smaller firms and these firms are more likely to conduct R&D in the same location as their headquarters.

Perhaps less obvious, if no less controversial, is how to define a firm's location in technological space. The Jaffe (1986) method (see section 2) requires detailed patent data for each firm, making it poorly suited for use with commonly available datasets such as COMPUSTAT. A less data-intensive method, albeit much less sophisticated, is to simply use each firm's SIC code.[25] COMPUSTAT assigns a single SIC to each firm by analyzing establishment industrial activity provided in firm *10K* reports and selecting the code considered to represent the firms major area of activity.[26] Thus, the measure of technological proximity used here is admittedly ordinal in that firms in same SIC group are defined as technologically near and all others as technologically distant. The technological proximity hypothesis is first examined assuming the technologically near spillover pool is limited to one's own four-digit SIC. This analysis is repeated at the three-digit SIC to examine the sensitivity of results to a broader definition of technological nearness.

Use of a firm's primary SIC to characterize technological location is far from ideal. A firm's true technological position is most certainly a multidimensional measure

[25] "[The SIC] was developed for use in the classification of establishments by type of activity … Each establishment is to be classified according to its primary activity." *Standard Industrial Classification Manual: 1987*, p. 11.
[26] *COMPUSTAT User's Guide* (1997), p. 249.

reflecting their diversity of innovative activity[27] or the diversity of their production operations or both. However, measurement error associated with use of the binary definition of technological proximity proposed here should impose a downward bias on parameter estimates. Thus, any evidence in their favor is arguably a robust finding.[28]

The raw financial variables are retrieved from the COMPUSTAT industrial, full coverage, and research annual data files. Observations in the raw data panel extend from 1970 to 1998[29]. Observations are selected for those firms whose primary industrial activity is within SIC 35 and whose corporate headquarters are within the 50 United States. Observations are deleted if the raw sales, employment, capital investment, or R&D expenditure entries are missing or combined with other variables. Further data trimming excludes observations where sales, employment, constructed capital stock, or constructed R&D stock are non-positive, leaving 4,680 observations.

An additional 105 observations are excluded which represent the lowest two percent of observations ordered by sales. This censoring is imposed to omit young, often high-tech firms from the panel that "go public" with little or no sales and are poor candidates for modeling with the production function framework. Indeed, 87 of the 105 excluded observations come from the highest tech sector in the panel; SIC 357 - Computer and Office Equipment. Preliminary analyses to check the sensitivity of regression estimates to this censoring indicate that the only result is to improve the goodness of fit of the own R&D coefficient. All other coefficient estimates are unaffected.

The final, unbalanced data panel includes 4,575 observations on 515 firms extending from 1972 to 1995. These data span 39 states and 29 four-digit (or eight three-digit) SIC's. Summary statistics are presented in table 3.

---

[27] Perhaps most properly computed in Jaffe (1986).
[28] Furthermore, Griliches (1992, p. S33) suggests that "[o]ne could argue that this is what the SIC classification is for. Presumably, the usefulness of somebody else's research to you is highest if he is in the same four-digit SIC classification as you are; it is still high if he is in the same three-digit industry group; and, while lower than before, the results of research by a firm in your own two-digit classification (but not three-digit) are more likely to be valuable to you than the average results of research outside it."
[29] Although the annual financial variables extend back to 1950, COMPUSTAT has only been tabulating R&D expenditures since 1970.

| **Table 3**<br>**Final Data Panel Summary Statistics; n = 4,575** | | | | |
|---|---|---|---|---|
| Variable | mean | std. dev. | min | max |
| sales ($M) | 619 | 2,100 | 0.299 | 41,400 |
| capital stock ($M) | 294 | 1,040 | 0.010 | 15,100 |
| employment (k) | 7.71 | 22.5 | 0.003 | 438 |
| R&D stock ($M) | 143 | 607 | 0.001 | 10,100 |

Spillover pools are seeded with a value of one in order to ensure they are well defined under the log transformation for all firm observations. This is a problem with the geographically near/technologically near spillover pool variable in particular which will be empty for firms that are in a remote location from others in their industry. Table 4 reports the average spillover pool size by total R&D and number of firms in each pool assuming a geographically near spillover pool boundary radius of 50 miles and a technologically near spillover pool defined by a firm's four-digit SIC.

| **Table 4**<br>**Spillover Pool Summary:**<br>**average pool R&D stock / average number firms in pool** | | |
|---|---|---|
| | geographically near (within 50 miles) | Geographically distant (beyond 50 miles) |
| technologically near (in same 4-digit SIC): | 84/0.6 | 857/9.2 |
| technologically distant (outside own 4-digit SIC): | 1,500/9.0 | 26,000/217 |

Recall that a principal motivation of this research is to assess the impact of industrial agglomeration on our interpretation of evidence of inter-firm spillovers. It is presumed that firms in the same industry tend to cluster spatially which, to some degree, will make evidence of technologically mediated spillovers and geographically mediated spillovers observationally equivalent. Before we begin the regression analysis, therefore, it is interesting to consider the extent to which industrial agglomeration is present in the particular data panel used in this study.

Table 5 presents a comparison of the probability that any pair of firms in the panel are in the same industry (column b) to the conditional probability that they are in the same industry given they are in the 'same location' (column c). Firms are considered to be in the 'same location' if the distance between them does not exceed the specified value (column a). Firms are considered to be in the same industry if they share the same four-digit SIC. Conditional probabilities are calculated assuming 'same location' is defined with a 25, 50, and 100 mile radius around each firm *i*.

**Table 5**
**Industrial Agglomeration among 515 Firms in the Panel**

| (a) | (b) | (c) | (d) | (e) |
|---|---|---|---|---|
| Maximum distance for any pair of firms *i* and *j* to be considered in the 'same location': | $Pr[SIC_j=SIC_i]$ | $Pr[SIC_j=SIC_i \mid loc_j=loc_i]$ | $d_m$ = (c) - (b) | t-$H_0$: $d_m=0$ |
| 25 miles | 0.052 | 0.289 | 0.237 | |
| | (0.022) | (0.158) | (0.158) | 34.0 |
| 50 miles | 0.052 | 0.217 | 0.165 | |
| | (0.022) | (0.149) | (0.150) | 25.0 |
| 100 miles | 0.052 | 0.172 | 0.120 | |
| | (0.022) | (0.132) | (0.132) | 20.6 |

* standard errors in parentheses.

Column (b) indicates that for each firm *i*, an average of 5.2 percent of the remaining firms in the panel are in the same four-digit SIC. Column (c) reports these values conditional on the definition provided in (a). For example, if we consider only the subset of firms within 50 miles of each other, fully 21.7 percent are in the same narrowly defined industry group. That is, firms within 50 miles of one another are over four times as likely to be in the same four-digit SIC than are any pair of firms from the panel in general.

The mean difference between the baseline (column b) and conditional (column c) probabilities for each firm is computed (column d) to test the hypothesis that these samples are selected from the same population. Returning to our example, the probability that any firm within 50 miles of another is in the same industry is, on average, 16.5 percentage points higher than the baseline probability. This sample mean is highly statistically significant, revealing evidence of industrial agglomeration in these data and suggesting inter-firm measures of geographic proximity will be correlated with measures of technological proximity.

Equation 7 presents the restricted form of the model derived in section 3 in an error components framework.[30]

$$y_{it} = \Gamma_0' D_{it} + \gamma_1 c_{it} + \gamma_2 l_{it} + \gamma_3 k_{it} + \gamma_4 \sum_{n \neq i}^{G_{Ni}} k_{nt-1} + \gamma_5 \sum_{z \neq i}^{G_{Zi}} k_{zt-1} + (u_i + \varepsilon_{it}) \qquad (7)$$

The model is restricted in the sense that external R&D stocks will be sorted into only two of four possible spillover pools. $D_{it}$ is a vector of location, sector, and time-specific dummy variables for firm *i* in time *t*. State and four-digit SIC dummies are included to control for a broad range of technological opportunity shocks and other narrowly defined fixed-effects. Individual year dummies are included to control for non-linear time specific technological opportunities.

Section 5 begins by considering the hypothesis that R&D spillovers are greatest from those firms that are geographically proximate, defined as those firms within 50 miles of the spillover receiving firm. For each firm *i*, all other firms are sorted into either a geographically near pool, $G_{Ni}$, or geographically distant pool, $G_{Zi}$. Each spillover pool variable is constructed by summing over the natural log transformed R&D stocks of all the firms in the respective pool.

Capital, labor, and own R&D stocks are assumed to become productive contemporaneously with own output. Spillover pool stocks are assumed to become productive at a one-year lag to reflect the additional time it takes to internalize publicly available knowledge.[31] Evenson (1968) examines aggregate data for U.S. agriculture and concludes that the lag structure of R&D takes an inverted V shape. He concludes that the peak weight from R&D flows is at five to eight year lags and little contribution is received from R&D expenditure at lags in excess of 10 to 16 years. But Wagner (1968) provides survey evidence that these lags are much shorter for industrial R&D, perhaps reflecting the more applied nature of private R&D expenditures. Assuming a constant

[30] Mairesse (1993, p. 433).
[31] A preliminary sensitivity analysis of this lag structure suggests the results reported below are robust. Various lags of own capital, own R&D, and R&D pools were considered. The main result seems to be to affect the point estimates of own input coefficient. The magnitudes and, more importantly, relative magnitudes of spillover pool coefficients are unaffected by this sensitivity.

rate of R&D expenditure, the 15 percent depreciation rate assumed in this study corresponds to an average R&D stock vintage of six years.

Due to the serially correlated nature of R&D time series, selecting the proper lag structure for R&D stock is a commonly recognized problem in the R&D productivity literature. As is generally the case, point estimates of R&D pool coefficients reported in the following sections are sensitive to alternative lag assumptions. For obvious reasons, however, this sensitivity is restricted to the time-series dimension. Furthermore, the qualitative findings of the analysis are unaffected by varying assumptions on lag length.

Following Mairesse (1993), between- (8.a) and within-firm (8.b) regressions are estimated in place of the total panel model presented in (7). This approach allows explicit examination of the cross-sectional and time-series variation in the data.

$$\bar{y}_i = \Gamma_0^{b}{}'\bar{D}_i + \gamma_1^b \bar{c}_i + \gamma_2^b \bar{l}_i + \gamma_3^b \bar{k}_i + \gamma_4^b \overline{\sum_{n \neq i}^{G_{Ni}} k_n} + \gamma_5^b \overline{\sum_{z \neq i}^{G_{Zi}} k_z} + (u_i + \bar{\varepsilon}_i) \tag{8.a}$$

$$\begin{aligned} y_{it} - \bar{y}_i &= \Gamma_0^{w}{}'(D_{it} - \bar{D}_i) + \gamma_1^w (c_{it} - \bar{c}_i) + \gamma_2^w (l_{it} - \bar{l}_i) + \gamma_3^w (k_{it} - \bar{k}_i) \\ &+ \gamma_4^w \left( \sum_{n \neq i}^{G_{Ni}} k_{nt-1} - \overline{\sum_{n \neq i}^{G_{Ni}} k_n} \right) + \gamma_5^w \left( \sum_{z \neq i}^{G_{Zi}} k_{zt-1} - \overline{\sum_{z \neq i}^{G_{Zi}} k_z} \right) + (\varepsilon_{it} - \bar{\varepsilon}_i) \end{aligned} \tag{8.b}$$

Following Bartlesman et al. (1994), between-firm variation coefficient estimates are interpreted as long-run elasticities as they reflect the response of output to permanent changes in variable values.

Estimates from the between-firm regression are often dismissed by econometricians as biased since the individual error term, $u_i$, is presumably correlated with the explanatory variables.[32] Estimates in 8.a will be biased to the extent that industry- and state-specific fixed effects do not control for firm specific variation in external (market and technological opportunities) and internal (personnel quality) fixed effects. Within-firm variation coefficient estimates are interpreted as short-run elasticities since they are driven by more transient changes in variable values. Mairesse

[32] Mairesse (1993, p. 427). For arguments in favor of within estimates due to biases expected in between estimates, see Mundlak (1961), Hoch (1962), and Mundlak (1978).

(1993) argues that these estimates may also be biased to the extent that time dummies do not control for temporal variation in capacity utilization.[33] Therefore, both between- and within-firm regression estimates are reported in the following chapter.

A two stage least squares (2SLS) procedure is used to control for estimation bias due to simultaneous choice of capital, labor, and output. This estimation procedure employs lagged values of own variable inputs as instruments for capital, labor, and own R&D. The regressions proposed in equations 8.a and 8.b are repeated to test the technological proximity hypothesis and finally the combined geographic/technological proximity hypothesis.

## 5. THE RESTRICTED ANALYSES: TESTING THE GEOGRAPHIC AND TECHNOLOGICAL PROXIMITY HYPOTHESES INDEPENDENTLY

This section presents between- and within-firm regression estimates from the restricted model as a point of departure for the general results and sensitivity analyses presented in the following section. These results illustrate that the framework proposed here can obtain findings consistent with the conventional wisdom that R&D spillovers are enhanced by both geographic and technological proximity.[34]

Table 6 reports regression estimates for equations 8.a and 8.b where "near" is defined as "within a 50 mile radius." All firms 50 miles or nearer to each firm $i$, regardless of their technological proximity, are pooled into the set $G_{Ni}$. The R&D of these firms compose the near spillover pool for firm $i$. The remaining set of firms, set $G_{Zi}$, define the distant spillover pool for firm $i$.

Goodness of fit is high in both the between- and within-firm regressions. The cross-sectional model exhibits nearly constant returns to scale with respect to conventional inputs, though the capital coefficient is not statistically significant. In the time-series regression, all own input coefficient estimates are significant at the five-

[33] Ibid. (p. 429). For arguments in favor of between estimates due to biases expected in within estimates, see Mairesse (1978) and Griliches and Mairesse (1984).
[34] The following analysis was initially conducted including the 2nd order terms derived in equation (4). However, these coefficient estimates were not significant. Also, see footnote 13.

**Table 6**
**Geographic Proximity Hypothesis: estimates assuming spillover pool boundary = 50 miles**[a]

| | Between[b] | Within[c] |
|---|---|---|
| *c* | 0.040 | 0.116** |
| | (0.035) | (0.022) |
| *l* | 0.975** | 1.026** |
| | (0.040) | (0.021) |
| *k* | 0.027** | 0.068** |
| | (0.011) | (0.010) |
| $\Sigma k_{Gn}$ | 0.012** | 0.005** |
| | (0.002) | (0.001) |
| $\Sigma k_{Gz}$ | 0.004** | 0.000 |
| | (0.001) | (0.001) |
| adj $R^2$ | 0.96 | 0.69 |
| n | 515 | 4,575 |

[a] standard errors in parentheses.
[b] intercept, state, and industry specific fixed effect coefficients not reported.
[c] intercept and year specific fixed effect coefficient estimates not reported.
* significant at the 10 percent level.
**significant at the 5 percent level.

percent level. The time-series model exhibits significant increasing returns to scale with respect to conventional inputs which is likely a result of measurement bias due to variation in capacity utilization over the business cycle.[35]

Turning to the spillover pool coefficients, the geographically near pool ($\Sigma k_{Gn}$) coefficient is significant in both the between- and within-firm regressions. In the between case, own output is three times more responsive to changes in R&D stocks less than 50 miles away than it is to changes in R&D beyond 50 miles. The near pool coefficient is larger than that of the distant spillover pool ($\Sigma k_{Gz}$) in the within case as well, a finding consistent with the claim the geographic distance attenuates inter-firm spillovers.

[35] Since the production function is primarily a technical relation, the correct measurement concept for independent variables is 'utilized inputs.' If capital, labor, or R&D stocks are costly to adjust then a firm will choose to overutilize (underutilize) inputs in the early stages of an economic expansion (contraction). Measurement error is introduced by use of purchased inputs as a proxy for utilized inputs. For example, consider a demand shock for a firm's product. Since capital is costly to adjust, the firm will increase both purchased inputs and utilization of existing stocks to service the increased demand. Total capital utilization will increase in proportion to output. But since input purchases understate this increase in utilization, the relatively large increase in output will be attributed to the relatively small increase in purchased inputs, imposing an upward bias on the capital coefficient estimate.

It is also interesting to compare the magnitude of the geographically near spillover pool coefficient to that of own R&D. Consider first the time-series dimension. A one percent shock to own R&D will result in an increase in own output of 0.068 percent in the short run. The impact of a change in R&D by the average firm in the nearby spillover pool is roughly 1/14 this magnitude. Recall from table 4 that, on average, approximately 10 firms are in the geographically nearby spillover pool. Therefore, a synchronous increase of R&D by firms in one's near spillover pool will result in an increase in output that is nearly as significant as the excess returns attributable to own R&D.[36]

Changes to external R&D appear even more important when we consider estimates obtained from the between-firm regression. In this case, changes in external R&D are nearly half as important as changes in own R&D. In the long run, a synchronous change in R&D of nearby firms may increase output by a magnitude greater than that of a comparable change in own R&D. This may seem an extreme finding but it is consistent with the view that, over time, returns to knowledge generating activities are difficult to contain.

Table 7 reports between- and within-firm regression spillover pool coefficient estimates as the definition of the nearby spillover pool is varied from 25 miles to 800 miles.[37] The geographically near pool coefficient exceeds that of the distant pool in nearly all cases. This difference, however, appears to diminish as the geographically near/distant spillover pool boundary is expanded. As the distance to the geographically near spillover pool boundary is increased, firms from the distant pool are shifted into the near pool. These firms are 'least near' relative to the composition of firms in the near spillover pool. This reallocation also removes the 'least distant' subset of firms from the

[36] The own R&D coefficient estimates presented here are understated due to failure to control for double counting of R&D expenditures. Schankerman (1981) showed that since these charges are also included in capital and labor variables, conventional input coefficient estimates account for normal returns to R&D inputs. The coefficient presented here is typically given an "excess returns interpretation." That is, the coefficient is interpreted to reflect only the risk premium associated with R&D expenditures rather than the total return. Schankerman's estimates of total returns to private R&D are up to three times larger than the excess returns estimates provided in table 2.3. Nevertheless, the magnitudes of spillover pool elasticities of output presented here appear significant and important.

[37] Conventional input coefficients and goodness of fit statistics are not reported since they are largely unchanged as compared with those reported in table 6.

distant pool. In both cases, external R&D stock coefficients will be biased downward if geographic distance attenuates spillovers.

**Table 7**
**Geographic Proximity Hypothesis: sensitivity of R&D coefficients to varying spillover pool boundary**[a]

| | miles: | 25 | 50 | 100 | 200 | 400 | 800 |
|---|---|---|---|---|---|---|---|
| BETWEEN | $k$ | 0.028** | 0.027** | 0.030** | 0.027** | 0.029** | 0.026** |
| | | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| | $\Sigma k_{Gn}$ | 0.015** | 0.012** | 0.009** | 0.006** | 0.007** | 0.005** |
| | | (0.003) | (0.002) | (0.002) | (0.001) | (0.002) | (0.001) |
| | $\Sigma k_{Gz}$ | 0.004** | 0.004** | 0.004** | 0.004** | 0.004** | 0.004** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| WITHIN | $k$ | 0.067** | 0.068** | 0.069** | 0.069** | 0.070** | 0.067** |
| | | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| | $\Sigma k_{Gn}$ | 0.008** | 0.005** | 0.002* | -0.000 | 0.001** | 0.001* |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{Gz}$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | -0.000 |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |

[a] standard errors in parentheses.
* significant at the 10 percent level.
**significant at the 5 percent level.

This sensitivity condition is satisfied for the geographically near spillover pool coefficient in both the cross-section and time series dimensions. In the between case, the nearby spillover pool coefficient drops by two thirds when the near/distant boundary is expanded from 25 to 800 miles. This reduction in coefficient value is even more extreme in the within case. Indeed, the nearby pool coefficient appears unimportant for distances in excess of 100 miles. The geographically distant spillover pool coefficients, however, do not appear to be diminished in the between case as this pool is restricted to only the most distant firms.

Comparison of nearby and distant spillover pool coefficients are consistent with the hypothesis that geographic proximity enhances spillover transmission. The output response to an increase in R&D from an average firm within 50 miles is three times as large as that for an average firm beyond this radius. This relation obtains in the time-series as well where distant spillover pool coefficients are insignificant. Furthermore, sensitivities of spillover pool coefficients to variation in the distance to the spillover pool boundaries are also consistent with the hypothesis that geographic distance attenuates R&D spillovers.

Table 8 reports regression estimates for equations 8.a and 8.b where "near" is defined as "within the same four-digit SIC." All firms within one's own narrowly defined industry group, regardless of geographic proximity, are pooled into the set $T_{Ni}$. The R&D of these firms compose the near spillover pool for firm *i*. The remaining set of firms, $T_{Zi}$, define the distant spillover pool for firm *i*.

**Table 8**

**Technological Proximity Hypothesis: estimates assuming spillover pool boundary = 4-digit SIC**[a]

| | Between[b] | Within[c] |
|---|---|---|
| *c* | 0.068** | 0.109** |
| | (0.034) | (0.021) |
| *l* | 0.953** | 1.033** |
| | (0.038) | (0.021) |
| *k* | 0.036** | 0.048** |
| | (0.010) | (0.010) |
| $\Sigma k_{Tn}$ | 0.042** | 0.011** |
| | (0.003) | (0.001) |
| $\Sigma k_{Tz}$ | 0.002** | -0.000 |
| | (0.001) | (0.001) |
| adj $R^2$ | 0.96 | 0.70 |
| n | 515 | 4,575 |

[a] standard errors in parentheses.
[b] intercept, state, and industry specific fixed effect coefficients not reported.
[c] intercept and year specific fixed effect coefficient estimates not reported.
* significant at the 10 percent level.
**significant at the 5 percent level.

Goodness of fit is high in both the between- and within-firm regressions. All own input coefficients are significant at the five percent level in both cases. This reflects an increase in magnitude and significance for the own capital coefficient which was insignificant in the between case in table 6. The cross-sectional model exhibits slightly increasing returns to scale with respect to conventional inputs. The time-series model exhibits significantly increasing returns to scale. The latter finding is again interpreted to reflect estimation bias due to variation in capacity utilization over the business cycle.

Turning to the spillover pool coefficients, the technologically near spillover pool ($\Sigma k_{Tn}$) coefficient is significant in both the between- and within-firm regressions. In both cases this coefficient is significantly larger than that of the distant spillover pool ($\Sigma k_{Tz}$), a finding consistent with the hypothesis that technological closeness accentuates inter-firm spillovers.

It is also interesting to compare the magnitude of the technologically near spillover pool coefficient to that of own R&D. The time-series dimension regression suggests a one percent shock to own R&D will result in an increase in own output of 0.048 percent in the short run. The impact of a change in R&D by the average firm in the nearby spillover pool is nearly one-fourth this magnitude. Recall from table 4 that, on average, there are approximately 10 firms are in the technologically near spillover pool. Therefore, a synchronous increase of R&D by all firms in this pool will result in an increase in output that is larger than the response from an increase in one's own R&D. This estimate is also larger than that from the prior section regarding short-run response to changes in the R&D stocks of geographic neighbors.

Changes in external R&D appear even more important when we consider estimates obtained from the between-firm regression. In this case, changes in the R&D of technological neighbors are at least as important as changes in own R&D. While this estimate may appear extreme, it is important to keep in mind that the own R&D coefficient is biased downward due to double counting of R&D expenditures in labor and capital variables.[38] Nonetheless, this result is not inconsistent with earlier findings in the literature of economically important returns from external R&D stocks.[39] Furthermore, the finding of large and significant spillovers from technological neighbors is consistent with the view advanced by Cohen and Levinthal (1989) that the objective of much R&D investment is to develop a firm's absorptive capacity; i.e. the ability to make use of rivals' ideas.

[38] See footnote 35. In contrast, while the within coefficient is subject to the same downward bias, it is also subject to an upward bias associated with procyclical variation in capacity utilization.
[39] See footnote 5.

Table 9 reports spillover pool coefficient estimates from the between- and within-firm regressions where the definition of the nearby spillover pool is relaxed from the four-digit to the three-digit SIC level.[40] This sensitivity results in reallocating the most technologically similar firms from the distant spillover pool to the near spillover pool. If technological similarity accentuates spillovers the external R&D stock coefficients will be biased downward as the definition of the spillover pool boundary is expanded. This condition is satisfied for both the near and distant spillover pools in both the cross-sectional and time-series regressions. In the between-firm regression, the near pool coefficient is larger than that of the distant pool for technological pool boundaries defined at both the four-digit and the three-digit levels. Furthermore, as we move from the four-digit to the three-digit specification, effectively sweeping technologically dissimilar firms into the nearby pool, this coefficient is biased downward. The technologically distant spillover pool coefficient is also biased downward as this pool is increasingly limited to

**Table 9**
**Technological Proximity Hypothesis: sensitivity of R&D coefficients to varying spillover pool boundary[a]**

| | pool boundary: | 4-d SIC | 3-d SIC |
|---|---|---|---|
| BETWEEN | $k$ | 0.036** | 0.028** |
| | | (0.010) | (0.009) |
| | $\Sigma k_{Tn}$ | 0.042** | 0.013** |
| | | (0.003) | (0.001) |
| | $\Sigma k_{Tz}$ | 0.002** | -0.002* |
| | | (0.001) | (0.001) |
| WITHIN | $k$ | 0.048** | 0.055** |
| | | (0.010) | (0.009) |
| | $\Sigma k_{Tn}$ | 0.011** | 0.007** |
| | | (0.001) | (0.001) |
| | $\Sigma k_{Tz}$ | -0.000 | -0.001** |
| | | (0.001) | (0.000) |

[a] standard errors in parentheses.
* significant at the 10 percent level.
**significant at the 5 percent level.

[40] Conventional input coefficients and goodness of fit statistics are not reported since they are largely unchanged by this sensitivity

the most technologically dissimilar firms in the panel.

Time-series estimates are also consistent with the hypothesis that spillovers from technologically similar firms are greater than are those from technologically distant firms. Nearby pool coefficients are larger than those of the distant spillover pool. In addition, the within-firm estimates of the technologically near spillover pool coefficient are biased downward as the pool boundary is expanded.

The combined results from tables 6 through 9 suggest that both technological and geographic proximity accentuate R&D spillovers. The strength of spillovers does appear to vary, with stronger spillovers estimated from a firms technological neighbors as compared to those from geographic neighbors. Nearby spillover pool coefficients from the relatively weakest definition of technological closeness (Table 9, last column) are roughly equal to those from the most restrictive definition of geographic closeness employed in these analyses (Table 7, first column). Recall from the previous chapter, however, that evidence of industrial agglomeration is present in this data panel (Table 5). This finding suggests that inter-firm measures of geographic proximity will be correlated with measures of technological proximity. Consequently, industrial agglomeration makes interpretation of the results presented in this section problematic.

## 6. THE GENERAL ANALYSIS: TESTING THE GEOGRAPHIC AND TECHNOLOGICAL PROXIMITY HYPOTHESES SIMULTANEOUSLY

This section presents estimates from the general spillover model proposed in section 3 which simultaneously accounts for the roles of geographic and technological proximity. Equations 8.a and 8.b are respecified to contain four spillover pools that vary in both the geographic and technological dimensions.

Table 10 reports between- and within-firm regression estimates obtained by defining R&D stocks within 50 miles as geographically near and those within a firm's own four-digit SIC as technologically near. $\Sigma k_{GnTn}$, $\Sigma k_{GnTz}$, $\Sigma k_{GzTn}$, and $\Sigma k_{GzTz}$ represent R&D spillover pools that are geographically near / technologically near, geographically near / technologically distant, geographically distant / technologically near, and geographically distant / technologically distant, respectively.

**Table 10**
**Combined Geographic and Technological Proximity Hypothesis: geographic boundary = 50 miles; technological boundary = 4-digit SIC[a]**

| | Between[b] | Within[c] |
|---|---|---|
| *c* | 0.056* | 0.105** |
| | (0.033) | (0.021) |
| *l* | 0.963** | 1.030** |
| | (0.037) | (0.021) |
| *k* | 0.037** | 0.048** |
| | (0.010) | (0.010) |
| $\Sigma k_{GnTn}$ | 0.032** | 0.010** |
| | (0.007) | (0.004) |
| $\Sigma k_{GnTz}$ | 0.009** | 0.005** |
| | (0.002) | (0.001) |
| $\Sigma k_{GzTn}$ | 0.030** | 0.011** |
| | (0.003) | (0.001) |
| $\Sigma k_{GzTz}$ | 0.002** | -0.000 |
| | (0.001) | (0.001) |
| adj $R^2$ | 0.96 | 0.70 |
| n | 515 | 4,575 |

[a] standard errors in parentheses.
[b] intercept, state, and industry specific fixed effect coefficients not reported.
[c] intercept and year specific fixed effect coefficient estimates not reported.
* significant at the 10 percent level.
**significant at the 5 percent level.

Goodness of fit is high in both the between- and within-firm regressions. In the between case, capital and labor coefficients are significant at the 10 and five percent levels, respectively. Own R&D is also significant at the five percent level. The cross-sectional model exhibits constant or slightly increasing returns to scale with respect to conventional inputs. All own input coefficient estimates are significant at the five percent level in the time-series regression. Again, this model exhibits significant increasing returns to scale with respect to own factor inputs.

Turning to the spillover pool coefficients, it is now possible to isolate the effects of geographic proximity from technological proximity. First, consider the importance of technological proximity for spillovers *after* controlling for geography. To do so we compare coefficients from technologically near and distant R&D that meet the same geographic criteria. In the between case, consider spillovers from R&D stocks within 50 miles of the spillover receiving firm. The output response to R&D from the same four-digit SIC is over three times larger than that from outside of a firm's own four-digit SIC (0.032 vs. 0.009). This relation holds for geographically distant R&D stocks as well: spillovers from R&D beyond 50 miles are greatest from that within a firm's own four-digit SIC (0.030 vs. 0.002).

Qualitatively similar results are obtained from the within-firm regressions. Spillovers from geographically near R&D are twice as large from that within a firm's own four-digit industry. However, spillovers from geographically distant R&D are *only* observed from within a firm's technology group.

The independent role of geographic proximity can be assessed by comparing spillover pool coefficients that meet common technological criteria. In the between case, a comparison of $k_{GnTn}$ to $k_{GzTn}$ pool coefficients (0.032 v. 0.030) indicates that spillovers from within a firm's own four-digit SIC do not appear to be attenuated by distance. In contrast, our intuition regarding the importance of geographic proximity is supported if we consider only spillovers from technologically dissimilar firms. Within this subset, spillovers are greatest from geographic neighbors (0.009 vs. 0.002). This same pattern is observed from the within-firm regression results.

These results present a much richer characterization of R&D spillovers than that provided by the restricted analysis reported in the previous section. The general analysis results are consistent with evidence of the importance of technological proximity obtained in the restricted analysis. Technological proximity appears to enhance inter-firm spillovers regardless of geographic proximity. But evidence regarding the importance of geographic proximity obtained from the general analysis is mixed. Geographic distance does attenuate spillovers between technologically dissimilar firms. However, spillovers among firms in the same four-digit SIC do not appear attenuated by distance. Thus, at least part of the correlation between output and the R&D of geographic neighbors appears to be driven by the tendency of technologically similar firms to agglomerate in space.

As in the previous section, additional evidence regarding the role of geographic and technological factors for R&D spillovers can be obtained by varying the spillover pool boundaries. Tables 11 and 12 report spillover pool coefficient estimates for the between- and within-firm regressions, respectively. The 'geographically near' pool specification is varied from 50 to 800 miles and the technologically near specification is

**Table 11**

**Between-firm Regression Estimates: combined geographic and technological proximity hypothesis**[a]

| | miles: | 50 | 100 | 200 | 400 | 800 |
|---|---|---|---|---|---|---|
| 4-d. SIC | $k$ | $0.037^{**}$ | $0.038^{**}$ | $0.036^{**}$ | $0.038^{**}$ | $0.036^{**}$ |
| | | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| | $\Sigma k_{GnTn}$ | $0.032^{**}$ | $0.038^{**}$ | $0.032^{**}$ | $0.034^{**}$ | $0.033^{**}$ |
| | | (0.007) | (0.007) | (0.006) | (0.004) | (0.004) |
| | $\Sigma k_{GnTz}$ | $0.009^{**}$ | $0.005^{**}$ | $0.002^{*}$ | $0.004^{**}$ | $0.002^{*}$ |
| | | (0.002) | (0.002) | (0.001) | (0.002) | (0.001) |
| | $\Sigma k_{GzTn}$ | $0.030^{**}$ | $0.030^{**}$ | $0.031^{**}$ | $0.030^{**}$ | $0.031^{**}$ |
| | | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) |
| | $\Sigma k_{GzTz}$ | $0.002^{**}$ | $0.002^{**}$ | $0.002^{**}$ | $0.002^{**}$ | $0.002^{**}$ |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| 3-d. SIC | $k$ | $0.029^{**}$ | $0.028^{**}$ | $0.027^{**}$ | $0.029^{**}$ | $0.027^{**}$ |
| | | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| | $\Sigma k_{GnTn}$ | $0.022^{**}$ | $0.021^{**}$ | $0.016^{**}$ | $0.016^{**}$ | $0.014^{**}$ |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) |
| | $\Sigma k_{GnTz}$ | 0.001 | -0.001 | $-0.004^{**}$ | -0.001 | $-0.002^{*}$ |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.001) |
| | $\Sigma k_{GzTn}$ | $0.012^{**}$ | $0.012^{**}$ | $0.012^{**}$ | $0.012^{**}$ | $0.012^{**}$ |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{GzTz}$ | $-0.002^{**}$ | $-0.002^{**}$ | $-0.002^{**}$ | $-0.002^{*}$ | -0.001 |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |

[a] standard errors in parentheses.
* significant at the 10 percent level.
**significant at the 5 percent level.

relaxed from the four-digit to the three-digit SIC.[41]

In the between case, 'technologically near' spillover pools coefficients ($\Sigma k_{GnTn}$ and $\Sigma k_{GzTn}$) are largely unchanged as the near pool radius is expanded from 50 to 800 miles. The only possible exception to this rule is observed in the $\Sigma k_{GnTn}$ coefficient in the lower portion of table 11. This pool is composed of R&D within the specified distance of the spillover receiving firm and in the same three-digit SIC. Under this broader definition of technological closeness the pool coefficient becomes relatively less important as distance is increased (i.e. relative to the corresponding row in the upper portion of the table). This finding suggests that spillovers from within a firms own three-digit SIC but outside its four-digit SIC may be sensitive to inter-firm distance.

The evidence form this claim is even more compelling when we consider the behavior of the $\Sigma k_{GnTz}$ coefficient as we vary the definition of technological closeness. In the bottom portion of table 11, this pool has been constrained to include only geographically near firms outside the broader three-digit SIC. By comparison to the respective row from the upper portion of the table, it appears that evidence of the importance of geographic proximity is driven by the subset of R&D that is technologically close (within same three-digit SIC) but not too close (within same four – digit SIC). Geographic distance has little attenuating effect on spillovers among very technologically similar firms.

The geographically near/technologically distant spillover pool coefficient behaves in a more intuitively appealing way. This coefficient suggests that spillovers from R&D stocks within a firm's own narrowly defined, four-digit SIC but within 50 miles are significant. Furthermore, as the definition of geographic proximity is relaxed to include more distant R&D stocks into this pool the coefficient is attenuated. These findings provide a richer characterization of the R&D spillover process, suggesting a departure between spillovers from other firms within one's own SIC and those from firms without.

[41] Conventional input coefficients and goodness of fit statistics are not reported since they are largely unchanged by this sensitivity It should also be noted that this sensitivity does not extend down to the 25 mile radius since, by construction, such a narrowly defined spillover pool is empty for many observations.

The lower half of table 11 relaxes the technological boundary of the near pool to include firms in the broader, three-digit SIC. As expected, all of the spillover pool coefficient estimates are biased downward by an expansion of the technological boundary from the four-digit to the three-digit level. As alluded to earlier, the behavior of the $\Sigma k_{GnTz}$ pool coefficient provides the most interesting result. This coefficient, the only one to indicate spillover attenuation as the geographic radius was expanded in the upper panel, becomes insignificant in all cases. This estimate suggests that no spillovers are received from very technologically dissimilar firms (those outside one's own three-digit SIC), even from those most geographically proximate.

It should be reiterated that some evidence of the importance of geographic proximity emerges in the $\Sigma k_{GnTn}$ pool as the technological boundary is expanded to the three-digit SIC. The $\Sigma k_{GnTn}$ pool coefficient, while insensitive to expansion of pool boundaries in the top half of table 11, appears to be declining, even if only slightly, in the bottom half of the table. Given the measurement problems in our proxy for firm location, this result would at least recommend additional research into the localization of spillovers among firms in the same three-digit SIC. The bias reported in table 11 is consistent with the claim that spillovers from R&D outside of a firms own four-digit SIC but within one's own three-digit SIC are attenuated by distance.

The within-firm regression results are qualitatively similar to those obtained in the cross section. However, a few exceptions can be observed. First, the $\Sigma k_{GnTn}$ coefficient appears to be biased downward as the geographic boundary is expanded in the top panel of table 12, though not in the bottom panel. Another difference from the between results is that the $\Sigma k_{GnTn}$ coefficient now appears to be biased downward as the geographic boundary is expanded in the top panel of table 12 rather than in the bottom panel. That is, in the short-run, spillovers from R&D in a firm's narrowly defined four-digit SIC are sensitive, if only slightly so, to geographic distance.[42] The $\Sigma k_{GnTz}$ coefficient in the top panel still suggests some evidence of the importance of technological proximity. And

[42] Although, the imprecision of the 50 and 100 mile radius pool coefficient estimates makes interpretation difficult.

again, this coefficient appears to reflect the importance of the R&D within a firm's own three-digit SIC, as the coefficient becomes insignificant in the lower panel of table 12.

**Table 12**
**Within-firm Regression Estimates: combined geographic and technological proximity hypothesis**[a]

| | miles: | 50 | 100 | 200 | 400 | 800 |
|---|---|---|---|---|---|---|
| 4-d. SIC | $k$ | 0.048** | 0.048** | 0.049** | 0.050** | 0.047** |
| | | (0.010) | (0.010) | (0.010) | (0.010) | (0.010) |
| | $\Sigma k_{GnTn}$ | 0.010** | 0.009** | 0.004* | 0.004** | 0.005** |
| | | (0.004) | (0.003) | (0.002) | (0.002) | (0.001) |
| | $\Sigma k_{GnTz}$ | 0.005** | 0.002** | -0.000 | 0.001** | 0.001** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{GzTn}$ | 0.011** | 0.011** | 0.012** | 0.014** | 0.016** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{GzTz}$ | -0.000 | -0.000 | -0.000 | -0.000 | -0.001* |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| 3-d. SIC | $k$ | 0.056** | 0.058** | 0.057** | 0.058** | 0.053** |
| | | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| | $\Sigma k_{GnTn}$ | 0.006** | 0.004** | 0.003** | 0.004** | 0.005** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{GnTz}$ | 0.001 | 0.001 | -0.001 | 0.000 | 0.000 |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{GzTn}$ | 0.007** | 0.007** | 0.007** | 0.007** | 0.007** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| | $\Sigma k_{GzTz}$ | -0.001** | -0.001** | -0.001** | -0.002** | -0.003** |
| | | (0.000) | (0.000) | (0.001) | (0.000) | (0.001) |

[a] standard errors in parentheses.
* significant at the 10 percent level.
**significant at the 5 percent level.

Another departure from the between results may be observed in the relative sizes of the technologically near coefficients ($\Sigma k_{GnTn}$ and $\Sigma k_{GzTn}$) in the upper panel. The results indicate that, among technologically similar firms, geographically near R&D generates fewer spillovers than geographically distant R&D when the geographic boundary exceeds 200 miles. This result obviously begs for further research.

The conclusions from this general analysis reveal subtleties masked by the partial analyses provided in section 5. While geographic distance appears to play some role in attenuating R&D spillovers, the effect appears to be limited to spillovers from outside (but not too far outside) a firm's own narrowly defined, four-digit SIC. It does not appear to matter whether R&D in a firm's own four-digit SIC is conducted geographically near or distant. Spillovers from R&D outside a firm's own four-digit SIC do appear to be attenuated by distance. However, these spillovers appear to be driven by R&D within a spillover receiving firm's own three-digit SIC. No evidence of spillovers are observed from technologically dissimilar firms when this pool is restricted to contain only that R&D from outside a firm's own three-digit SIC.[43]

## 7. CONCLUDING DISCUSSION

This analysis improves our understanding of the importance of geographic and technological factors for inter-firm spillovers from R&D activity. Parameter estimates obtained in a production function framework indicate that spillovers are significant and important from geographically and technologically proximate R&D stocks. The restricted analysis results presented in section 5 are consistent with intuition and existing empirical evidence which suggests that both geographic and technological distance attenuate knowledge spillovers. Results from the general analysis reported in section 6, however, suggest that the importance of geographic proximity is conditional on the technical relation between spillover sending and receiving units. This study demonstrates that R&D spillover estimates obtained by methods that fail to control for industrial agglomeration may not reveal the conditional nature of the importance of geography.

In a preliminary analysis, "geographically near R&D" is defined as that conducted within a specified radius of each firm and is varied from 25 to 800 miles. The preliminary results support the conventional wisdom that the magnitude of spillovers depends on geographic proximity. Spillovers are most significant from R&D stocks within a 50 mile radius and their strength is attenuated with increasing distance. These conclusions, however, are found to be dependent on the propensity for industrial agglomeration. Results generated from the more general analysis that controls for

[43] It should be noted that some evidence of spillovers from outside the three-digit SIC is observed if a pool is constrained to contain only firms within 25 miles. However, these estimates are relatively imprecise since such a narrowly defined pool is empty by construction for large number of observations. Recall footnote 41.

correlation between measures of geographic and technological proximity reveal that evidence of the importance of geographic proximity is limited to spillovers from R&D outside a firm's own narrowly defined industry group.

The hypothesis that technological proximity generally magnifies spillovers is supported by these analyses. "Technologically near R&D" is defined as that conducted within a firm's own industry group. Industry group is defined at the four-digit and then the three-digit Standard Industrial Classification (SIC). The hypothesis is supported by the preliminary analysis in which external R&D stocks are sorted into technologically near and distant pools. The correlation between own output and R&D from within a firm's own four-digit SIC is positive and significant. This coefficient is diminished when technological proximity is defined at the broader three-digit level. This conclusion is supported by results obtained from the more general framework. Spillovers are greatest among technologically similar firms, regardless of geographic proximity, and are attenuated as technological distance increases.

Table 10 indicates that the short-run return from R&D stocks within a firm's own four-digit SIC is, on average, about one-fifth as important as changes in a firm's own R&D. A one percent increase in the R&D stock of a single firm in the same industry implies a 0.01 percent increase in own output. In the longer-run, these spillovers are nearly as important as the contribution of own R&D stocks.[44] These magnitudes are unaffected by geographic distance between spillover sending and receiving firms—the same average elasticity is reported for R&D within and beyond a 50 mile radius around each firm.

In contrast, returns from the R&D of technologically distant firms are sensitive to geographic proximity to the spillover receiver. In the short-run, the returns from the R&D of each technologically dissimilar firm within 50 miles are half as large as those from technologically similar firms (0.005 vs. 0.010). In the long-run the factor is three (0.032 vs. 0.009). But long-run returns from the R&D of technologically dissimilar firms

[44] Recall, however, that the own R&D coefficient is underestimated due to double counting with no countervailing bias, as in the within case, associated with procyclical variation in capacity utilization.

beyond 50 miles are less than one-fourth the magnitude of their geographically proximate counterparts. And short-run returns are insignificant altogether.

These findings are consistent with existing evidence of the importance of geographic and technological proximity for knowledge spillovers. While Jaffe et al. (1993) present compelling evidence that spillovers from inventive activity are localized, these results suggest that spillovers from innovative activity may behave differently. They find that patents have a higher likelihood of proximity to their own citations than to patents in a technologically similar control group. Their findings suggest that scientists who define the 'state of the art' benefit from proximity to potential collaborators. The evidence presented in this study, however, suggests that within-industry spillovers from industrial research activity that result in marketable innovations are not localized. To the extent that localization externalities from industrial R&D do exist they appear to be attributable to the application of knowledge from outside a firm's own narrowly defined technology group. This interpretation suggests that the presumably more applied set of activities summarized by private R&D expenditures result in spillovers that are quickly transmitted through firms in the same industry.[45]

The significance and magnitude of R&D spillovers documented here support continued research into the role of public policy intervention in the R&D arena. Positive externalities to innovative activity suggest that a decentralized system of investment may not provide sufficient incentives to generate the socially optimal level of R&D activity. However, several studies have documented the superior profitability of private R&D investment over its public counterpart, arguing against government funded research. While this return premium may be attributable to the more applied nature of private R&D expenditures, it also reflects a difference of incentives directing employment of resources in each of these sectors.

[45] If academic scientists and institutions are less likely to capitalize residual claims on inventions, one may also argue that the decision to collaborate is the outcome of a cost minimization calculation. In contrast, firms who may be more effective at appropriating rents from innovation may be more accurately modeled with a profit maximization objective and more likely to justify the cost of travel for the purpose of collaboration and knowledge transfer.

The geographic and technological pattern of spillovers documented here has implications for our assessment of the efficiency of concentrated market structures in knowledge intensive industries. It was argued in section 2 that allocative inefficiency associated with market power may be offset by an improvement in dynamic efficiency associated with internalization of knowledge spillovers. This argument is often offered in defense of large firms whose market dominance is challenged on antitrust grounds. The results presented in this study suggest that knowledge spillovers are largest among firms in the same narrowly defined industry. One may therefore offer a preliminary argument in defense of increased concentration in particular industries. Presumably, to the extent that larger firms internalize a larger fraction of total returns to innovative activity they will invest in more of it. Among technologically similar firms, the partial spillover enhancing effect of geographic proximity is much less significant. A defense of mergers between firms in a particular geographic region therefore may not be justified by the 'internalization of knowledge spillovers' argument.

The evidence presented in this study is particularly relevant to the current debate regarding the role of knowledge spillovers in city formation. These findings are consistent with those of Feldman and Audretsch (1999) who report "a tendency for innovative activity in complementary industries to cluster together in geographic space."[46] This evidence is consistent with the argument advanced by Jacobs (1969) that firms in urban agglomerations are more productive because they are able to make use of ideas from other industries. This is not to say MAR externalities, knowledge spillovers among firms in the same industry, are unimportant. Indeed, the estimates presented in chapter five indicate that they are more important and more likely to be statistically significant than are Jacobs externalities. However, MAR 'localization' externalities do not appear to serve a localizing function. These findings suggest that knowledge spillovers may provide the centripetal force necessary for industrially diverse agglomerations. However, the set of activities that compose industrially homogenous cities are likely drawn together by something other than geographically dependent

[46] Feldman and Audretsch (1999, p. 411).

knowledge spillovers. Other potential agglomerating forces such as natural advantages or shared intermediate inputs deserve increased attention in the empirical literature.

Finally, these results may serve to guide future empirical research related to the new growth theory and also have implications for understanding the process of economic development and the effectiveness of development policy. The puzzling empirical regularity of conditional convergence is contrary to the predictions of basic knowledge-spillover growth models. In these models, the public goods nature of knowledge suggests economies should share a common set of production possibilities. The foregoing analysis would suggest convergence clubs may be defined by similarity of members' technological bases, regardless of geographic proximity.[47] If knowledge is transferred among technologically similar firms then our understanding of conditional convergence may be served by an analysis of institutional arrangements that facilitate industrial convergence. In contrast, countries without moderately advanced technology intensive sectors may receive little benefits from knowledge spillovers from more highly developed economies.

[47] For a recent theoretical contribution along these lines, see Basu and Weil (1998).

## References


Adams, James and Adam Jaffe, 1996, "Bounding the Effects of R&D: an investigation using matched establishment firm data," *RAND Journal of Economics*, 27: 700-721.

Aghion, Phileppe and Peter Howitt, 1992, "A Model of Growth Through Creative Destruction," *Econometrica*, 60(2): 323-351.

Arrow, Kenneth J., 1962, "The Economic Implications of Learning by Doing," *Review of Economic Studies*, 29: 155-73.

Audretsch, David and Maryann Feldman, 1996, "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86(3): 630-640.

Bartlesman, Eric J., Ricardo J. Caballero, and Richard K. Lyons, 1994, "Customer- and Supplier-Driven Externalities," *American Economic Review*, 84(4): 1075-1084.

Basu, Susanto, 1996, "Procyclical Productivity: Increasing Returns or Cyclical Utilization?," *Quarterly Journal of Economics*, 111(3): 719-51.

____________, and David N. Weil, 1998, "Appropriate Technology and Growth," *Quarterly Journal of Economics*, 113(4): 1025-54.

Bernstein, Jeffrey I., and M. Ishaq Nadiri, 1989, "Research and Development and Intra-industry Spillovers: An Empirical Application of Dynamic Duality," *Review of Economic Studies*, 56: 249-69.

Bowker, R. R., Database Publishing Group, 1992, *Directory of American Research and Technology 1993: Organizations Active in Product Development for Business*, 27th edition, New Providence, New Jersey: R. R. Bowker.

Cameron, Gavin, 1996, "Innovation and Economic Growth," *Centre for Economic Performance* Discussion Paper no. 277.

Chirinko, Robert S., Steven M. Fazzari, and Andrew P. Meyer, 1999, "How Responsive is Business Capital Formation to its User Cost? An Exploration with Micro Data," *Journal of Public Economics*, 74: 53-80.

Cockburn, Iain and Zvi Griliches, 1988, "Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents," *American Economic Review*, 78(2): 419-423.

Cohen, Wesley M. and Daniel A. Levinthal, 1989, "Innovation and Learning: The Two Faces of R&D," *Economic Journal*, 99(3): 569-96.

Cooper, Christopher, 1998, “Oil Firms Still Rely on Corporate Spies to be Well Informed,” *Wall Street Journal*, December 7: A, p. 1, col. 4.

Evenson, R. 1968, “The contribution of agricultural research and extension to agricultural production,” Unpublished Ph.D. thesis, University of Chicago.

Feldman, Maryann P. and David B. Audretsch, 1999, “Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition,” *European Economic Review*, 43(2): 409-29.

Glaeser, Edward L., Hedi D. Kallal, Josè A. Scheinkman, and Andrei Shleifer, 1992, “Growth in Cities,” *Journal of Political Economy*, 100(6): 1126-1152.

Grabowski, H. G., 1968, “The Determinants of Industrial Research and Development: a study of the chemical, drug, and petroleum industries,” *Journal of Political Economy*, 72:292-306.

Griliches, Zvi, 1958, “Research Cost and Social Returns:  Hybrid corn and related innovations,” *Journal of Political Economy*, 66(5): 419-431.

___________, 1979, “Issues in Assessing the Contribution of R&D to Productivity Growth,” *Bell Journal of Economics*, 10(1): 92-116.

___________, 1992, “The Search for R&D Spillovers,” *Scandinavian Journal of Economics*, 94:s29-s47.

____________, and Jacques Mairesse, 1984, “Productivity and R&D at the Firm Level,” in Zvi Griliches, ed., *R&D, Patents and Productivity*, Chicago: Chicago University Press.

Grossman, Gene and Elhanan Helpman, 1991, *Innovation and Growth in the Global Economy*, Cambridge, MA: MIT Press.

Hall, Bronwyn and Jacques Mairesse, 1995, “Exploring the Relationship between R&D and Productivity in French Manufacturing Firms,” *Journal of Econometrics*, 65: 263-293.

Henderson, J. Vernon, and A. Kuncoro, and M. Turner, 1995, “Industrial Development and Cities,” *Journal of Political Economy*, 103: 1067-81.

Henderson, Rebecca and Iain Cockburn, 1996, “Scale, scope, and spillovers: the determinants of research productivity in drug discovery,” *The RAND Journal of Economics*, 27(1): 32-59.

Hoch, I., 1962, “Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data,” *Econometrica*, 30: 34-53.

Jacobs, Jane, 1969, *The Economy of Cities*, New York: Random House.

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

____________, Manuel Trajtenberg, and Rebecca Henderson, 1993, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent citations," *The Quarterly Journal of Economics*, August 1993: 577-598.

Jorgenson, Dale W., and Kevin J. Stiroh, 2000, "Raising the Speed Limit: U.S. Economic Growth n the Information Age," *Brookings Papers on Economic Activity*, 1.

Levin, Richard C., Alvin K. Klevorick, Richard R. Nelson, and Sidney G. Winter, 1987, "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity,* 3: 783-820.

Mairesse, Jacques, 1978, "New Estimates of Embodied and Disembodied Technical Progress," *Annales de l'INSEE*, 30-31: 681-791.

______________, 1993, "Time-Series and Cross-Sectional Estimates on Panel Data: why they are different and why should they be equal?," in G. S. Maddala (ed.) *The Econometrics of Panel Data*, Volume 2. Elger Reference Collection Series. *International Library of Critical Writings in Econometrics*, Volume 1. UK: Aldershot.

Minasian, J., 1962, "The Economics of Research and Development," in Richard R. Nelson (ed.) *The Rate and Direction of Inventive Activity*, NBER Special Conference Series, vol. 13, Princeton, NJ: Princeton University Press.

Marshall, Alfred, 1890, *Principles of Economics,* London: Macmillan.

Mundlak, Yair, 1961, "Empirical Production Function Free of Management Bias," *Journal of Farm Economics*, 43: 44-56.

_____________, 1978, "On the Pooling of Time-Series and Cross-Section Data," *Econometrica*, 46: 49-85.

Nadiri, M. Ishaq, 1993, "Innovations and Technological Spillovers," *National Bureau of Economic Research* Working Paper No. 4423.

National Science Foundation, 1998, *Research and Development in Industry: 1996*, Arlington, VA: NSF.

Office of Management and Budget, Executive Office of the President, 1987, *Standard Industrial Classification Manual*, Springfield, VA: National Technical Information Service.

Oliner, Stephen D., and Daniel E. Sichel, forthcoming, “The Resurgence of Growth in the Late 1990’s: Is Information Technology the Story?” *Journal of Economic Perspectives*.

Romer, Paul M., 1986, “Increasing Returns and Long-Run Growth,” *Journal of Political Economy,* 94: 1002-37.

____________, 1990, “Endogenous Technological Change,” *Journal of Political Economy*, 98(5,2): S71-101.

Salinger, M. and L. Summers, 1983, “Tax Reform and Corporate Investment,” in Martin Feldstein (ed.), *Behavioral Simulation Methods in Tax Policy Analysis,* Chicago: University of Chicago Press.

Saxenian, AnnaLee, 1994, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128,* Cambridge, Massachusetts: Harvard University Press.

Schankerman, Mark, 1981, “The Effects of Double-Counting and Expensing on the Measured Returns to R&D,” *Review of Economics and Statistics*, 63(3) 454-58.

Schumpeter, J., 1942, *Capitalism, Socialism, and Democracy*, New York: Harper.

Spence, A. Michael, 1984, “Cost Reduction, Competition, and Industry Performance,” *Econometrica*, 52:101-121.

Standard & Poor, 1997, *COMPUSTAT User’s Guide*, Englewood, CO: McGraw-Hill.

Terleckyj, N., 1958, “Factors Underlying Productivity: some empirical observations,” *Journal of the American Statistical Association*, 53:593.

Wagner, L. U., 1968, “Problems in estimating research and development investments and stock,” In *1968 Proceedings of the Business and Economic Statistics Section*, Washington, DC: ASA, 189-97.