Measuring R&D Spillovers: On the Importance of Geographic and Technological Proximity

Michael J. Orlando

July 2000
Last Revised: **September 2002**

RWP 02-06

Research Division
Federal Reserve Bank of Kansas City

Michael Orlando is an economist at the Federal Reserve Bank of Kansas City. Omissions or errors are his responsibility alone and persist in spite of the invaluable guidance and comments provided by Bruce Petersen, Marcus Berliant, Steven Fazzari, Josh Lerner, Douglass North, Robert Parks, the participants of the NBER Summer Institute - Productivity, Science, and R&D, and two anonymous referees. 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.

e-mail: michael.j.orlando@kc.frb.org

Abstract

Evidence is presented which suggest that an important measure of the apparent geographic localization of R&D spillovers may be an artifact of industrial agglomeration. A production function framework is used to examine the role of geographic and technological proximity for inter-firm spillovers from R&D. The largest spillovers are found to flow between firms in the same industry. However, spillovers within narrowly defined technological groups do not appear to be attenuated by distance. Geographic proximity does appear to attenuate spillovers that cross narrowly defined technological boundaries, suggesting these spillovers may play a role in the agglomeration of a diversity of industrial activity.

JEL classification: O3, R1, L6

Key words: R&D, spillovers, industrial agglomeration, geography, empirical studies.

1. INTRODUCTION

Identification of the importance of geographic and technological proximity for research and development (R&D) spillovers is complicated if firms in an industry agglomerate for reasons exclusive of localized inter-firm knowledge spillovers.[1] The purpose of this paper is to distinguish the importance of these two factors for R&D spillovers. Direct evidence of spillovers is obtained from a firm-level production function framework that includes geographically and technologically proximate R&D stocks. Spillovers are found to be largest among firms *within* the same narrowly defined industry. However, these spillovers are largely insensitive to inter-firm distance. While R&D spillovers *across* narrowly defined industry boundaries are smaller in magnitude, they appear to be attenuated by distance. Consequently, R&D spillovers may be a force contributing to the formation of industrially diverse agglomerations.

Knowledge spillovers will play a role in agglomeration if they are attenuated by distance. This mechanism will result in localization, or the agglomeration of similar industrial activity, if knowledge spillovers benefit only firms in the same industry. Urbanization, or the agglomeration of a diversity of industrial activity, will occur if knowledge spills across industry boundaries.

A significant body of empirical evidence establishes the importance of spillovers from R&D.[2] A number of studies examine the role that geographic and technological distance play in determining spillover intensity. Audretsch and Feldman (1996) find

---

[1] Among the firms considered in this study, for example, those within 100 miles of one another are over three times more likely to be in the same four-digit SIC group than are any two randomly chosen firms. And firms within 50 miles of one another are over four times more likely to come from the same four-digit SIC. While this degree of agglomeration may result if distance attenuates intra-industry spillovers, it may also result if firms in an industry benefit from a shared intermediate input or some natural advantage.

[2] Cameron (1996) provides a recent survey.

greater clustering of innovative activity in industries where knowledge externalities are likely to be relatively important, providing indirect evidence of a geographic bound to knowledge spillovers.[3] Jaffe (1986) finds that firms active in research intensive technology groups enjoy higher research productivity and higher returns to R&D, suggesting technological proximity accentuates spillovers. Adams and Jaffe (1996) consider both the importance of technological and geographic proximity for R&D spillovers. While they find evidence that spillovers are attenuated by both factors, the study focuses on spillovers at the intra-firm level. Moreover, data limitations do not permit the authors to consider the two hypotheses simultaneously.[4]

The next section discusses the data and extends the production function framework recommended by Hall and Mairesse (1995) to include external stocks of R&D.[5] Empirical results are summarized in section three. The implications of these findings are discussed in the concluding section.

2. EMPIRICAL SPECIFICATION AND DATA SELECTION

Hall and Mairesse (1995) is the point of departure, where firm output, Y_{it}, is represented with a conventional Cobb-Douglas production technology:[6]

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

where C_{it} is capital, L_{it} is labor, α and β are their respective output elasticities, and ε is an

[3] Jaffe et al. (1993) and Jaffe and Trajtenberg (1999) provide evidence of localization of knowledge spillovers from inventive activity.

[4] "Unfortunately, there are inherent limitations of the R&D data to identify the effects of distance along both geographic and technological dimensions. To estimate both effects one would need data revealing the *joint* distribution of research activity along the two dimensions. Instead we observe just the *marginal* distributions." Adams and Jaffe (1996) p. 703.

[5] See Orlando (2000) for a detailed exposition.

[6] 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.

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 at time *t*. Assuming R&D is the determinable component of knowledge relevant to the production process, then $\hat{K}_{it}$ represents all own and external stocks of R&D relevant to firm *i* production.

The empirical specification parameterizes the state of technology using a straightforward application of the knowledge production function (Pakes and Griliches 1984):

$$y_{it} = A_0' D_{it} + ac_{it} + bl_{it} + rk_{it} + s_{NN} \sum_{j \neq i}^{(G_N T_N)_i} k_{jt-1} + s_{NZ} \sum_{j \neq i}^{(G_N T_Z)_i} k_{jt-1} + s_{ZN} \sum_{j \neq i}^{(G_Z T_N)_i} k_{jt-1} + s_{ZZ} \sum_{j \neq i}^{(G_Z T_Z)_i} k_{jt-1} + (u_i + \varepsilon_{it}) \tag{2}$$

where lower case variables represent natural logarithms. External stocks of R&D are sorted into four spillover pools. $(G_N T_N)_i$ represents the subset of firms that are both geographically and technologically near to firm *i*. $(G_N T_Z)_i$ represents the subset of firms that are geographically near but technologically distant from firm *i*. $(G_Z T_N)_i$, and $(G_Z T_Z)_i$ are defined respectively. D_{it} is a vector of location, sector, and time-specific dummy variables, included to control for a broad range of technological opportunity shocks and other narrowly defined fixed-effects.[7]

Capital, labor, and own R&D stocks are assumed to influence output contemporaneously. Spillover pool stocks are assumed to influence production at a one-year lag to reflect the additional time it takes to internalize publicly available

---

[7] Given that firms share spillovers pools of the same size, the pool coefficients obtained in (2) imply that large firms receive a relatively larger benefit than small firms from the same size spillover pool. Since the non-rival nature of R&D is the central point of motivation, this assumption is intuitively appealing. Additional support is provided by Jaffe (1986) who reports benefits from R&D spillovers increasing in own scale of R&D and Henderson and Cockburn (1996) who conclude that larger research efforts in the pharmaceutical industry are more productive because they allow economies of scope to be realized.

knowledge.[8] Between-firm, within-firm, first-difference, and several long-difference regressions are estimated in place of the total panel model presented in (2). A two stage least squares (2SLS) procedure is used to control for simultaneity bias. Lagged values of own capital, labor, and R&D are used as instruments.

The present analysis considers firms in SIC 35, Industrial, Commercial Machinery, and Computer Equipment.[9] Annual firm-level financial data 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.

Firm sales is used to proxy for output. Estimation error associated with this proxy will be confined to the constant term if materials charges are a fixed proportion of sales and market power results in a constant mark-up. Basu (1996) reports that material inputs are nearly perfectly correlated with output.[10] While it is arguable that mark-up is constant across all observations in the panel, it may be so across firms in a particular industry. In this case, industry dummies will control for variation in market power. Assuming market power is constant across time for a given firm, the firm fixed effects regression models will control associated estimation error.

The capital stock variable is constructed by accumulating capital spending

---

[8] Evenson (1968) suggests the peak weight from R&D flows occurs at five to eight year lags and little contribution is received from R&D expenditures older than 10 to 16 years. Wagner (1968) concludes the lags are much shorter for industrial R&D, perhaps reflecting the more applied nature of private R&D expenditures. Assuming a constant 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. The results reported below are robust to a variety of lag structures.

[9] 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).

[10] In addition, Mairesse and Hall (1996) find that including materials directly into production function regressions can give misleading results and conclude that estimates based on sales are not badly biased.

following Salinger and Summers (1983). A correction for acquisitions and divestitures per Chirinko et al. (1999) avoids exclusion of firms engaged in mergers over the sample period.

The R&D stock variable is constructed from R&D expenditures using the perpetual inventory method commonly employed in studies of R&D productivity (Griliches 1979). The own R&D coefficient is interpreted as an "excess return" since data availability does not allow for correction of double-counting (Schankerman 1981). The results reported below are invariant to various depreciation and pre-sample period growth rates.[11]

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 technical professionals, reflecting the fact that the majority of R&D expenditures represent staff salaries.[12]

Firms are assumed to be located at the geographic centroid of the county location of their corporate headquarters. In the base case analysis, a 50 mile radius around each firm is used to define all other firms as geographically near (inside the circle) or

---

[11] Results reported below were insensitive to depreciation rates ranging from five to 30 percent. Hall & Mairesse (1995) also find coefficient estimates robust to the choice of depreciation rate in constructing R&D capital, a result that has been observed in a number of previous studies.

[12] 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.

geographically distant (outside the circle).[13] Location in technological space is defined using a firm's four-digit SIC.[14]

The COMPUSTAT data are taken from the industrial, full coverage, and research annual data files. The raw data panel includes firms with corporate headquarters within the 50 United States and extends from 1970 to 1998. Observations are deleted where 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, making them poor candidates for modeling with a production function framework.[15]

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 1.

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

---

[13] Robustness checks included at the end of the next section illustrate that the conclusions of the analysis are invariant to variations in the 'geographically near' radius and alternative measures of firm location.

[14] Currently, Jaffe (1986) provides the superlative index of relative technological position. Due to aggregation in the empirical strategy defined by (2), this study would not exploit the degree of variation inherent in such a measure. Measurement error associated with use of SIC's in this context should impose a downward bias on spillover pool coefficient estimates. Also, see Griliches (1992).

[15] Indeed, 87 of the 105 excluded observations come from the highest tech sector in the panel; SIC 357 - Computer and Office Equipment. The main result of this censoring is to improve goodness of fit of the own R&D coefficient.

Table 1 **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

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 2 reports the average spillover pool size by total R&D and number of firms in each pool.

Table 2 **Spillover Pool Summary:** **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

3. RESULTS

Table 3 presents results from between, within, and a range of difference regression estimates of equation (2). Goodness of fit is high in the between, within, and longer-differenced estimates. 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 regression suggests constant or slightly increasing returns to scale with respect to conventional inputs in the long-run. In the within and higher difference regressions, all own input coefficient estimates are significant at the five-percent level.[16] The time-series regressions exhibit significant increasing returns to scale with respect to conventional inputs, which is likely to result

Table 3
Base case regression results: geographically near ≤ 50 miles, technologically near = 4-d SIC[a]

Variable	Between[b]	Within[c]	1st diff.[d]	2nd diff.[d]	4th diff.[d]	8th diff.[d]
c	.056*	.105**	-6.28	.052	.076**	.080**
	(.033)	(.021)	(11.0)	(.036)	(.025)	(.031)
l	.963**	1.03**	12.4	1.15**	1.06**	1.06**
	(.037)	(.021)	(17.2)	(.035)	(.023)	(.030)
k	.037**	.048**	-.225	.043**	.041**	.029**
	(.010)	(.010)	(.590)	(.016)	(.011)	(.015)
$\sum k_{GnTn}$	.032**	.010**	.027	.003	.002	.013**
	(.007)	(.004)	(.060)	(.004)	(.004)	(.006)
$\sum k_{GnTz}$	.009**	.005**	-.007	-.000	.001	.005**
	(.002)	(.001)	(.015)	(.001)	(.001)	(.002)
$\sum k_{GzTn}$	.030**	.011**	-.004	.002**	.005**	.014**
	(.003)	(.001)	(.013)	(.001)	(.001)	(.002)
$\sum k_{GzTz}$	.002**	-.000	-.013	-.000	-.001	-.000
	(.001)	(.001)	(.021)	(.001)	(.001)	(.002)
adj R^2	.96	.70	-.00	.49	.71	.78
N	515	4,575	4,060	3,580	2,735	1,484

[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.
[d] year-specific fixed effect coefficient estimates not reported.
* significant at the 10 percent level.
**significant at the 5 percent level.

from measurement bias due to variation in capacity utilization over the business cycle.[17]

Turning to the spillover pool coefficients, the geographically and technologically near pool (Σk_{GnTn}) coefficient is significant in both the between- and within-firm regressions and returns to significance in the eighth difference.[18] It is larger than the

[16] Equation (2) was estimated through a 10th difference. All own-input coefficients are significant at the five-percent level in the 3rd through 10th difference regressions.

[17] 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.

[18] The $\sum k_{GnTn}$ pool coefficient is significant at the 10 percent level in the 7th difference regression and at the five-percent level in the 8th through 10th difference regressions. The $\sum k_{GnTz}$ pool coefficient is significant at

geographically near/technologically distant pool (Σk_{GnTz}) coefficient, suggesting R&D spillovers are stronger among firms in a narrowly defined industry. Comparing coefficients from the geographically distant spillover pools also supports this conclusion. Spillovers from this technologically near pool (Σk_{GzTn}) are significant in most cases and larger than those from the technologically distant pool (Σk_{GzTz}) in all cases.[19]

Comparison of coefficients from the technologically similar spillover pools suggests only spillovers across industry boundaries are attenuated by distance. In the between, within, and eighth difference regressions, the Σk_{GnTn} coefficient is similar in magnitude to the Σk_{GzTn} coefficient, suggesting spillovers between firms in the same, narrow SIC are not attenuated by distance. In contrast, the Σk_{GnTz} coefficient is significant and generally larger than the Σk_{GzTz} coefficient, indicating that spillovers received from outside a firm's own four-digit SIC are attenuated by distance.[20]

A wide range of sensitivities to the base-case were explored. Several of these results are included in Table 4 and discussed below. The eighth-difference regressions is used as a basis for comparison since the long-difference will control for correlated fixed effects while avoiding spurious, high-frequency correlation that may be the source of significance in the 'within' case. The base-case results are repeated in column (1) for comparison.

---

the five-percent level in the 7th through 10th difference regressions. The $\sum k_{GzTn}$ pool coefficient is significant at the five-percent level in the 2nd through 10th difference regressions.

[19] The importance of technological proximity for spillovers is suggested when equation (2) is estimated with the restrictions $s_{NN} = s_{ZN}$ and $s_{NZ} = s_{ZZ}$. In this case, the technologically near pool coefficient is significant and larger than the technologically distant pool coefficient.

[20] The distinction in geographic sensitivity of spillovers between technologically near and technologically distant firms is not identified if equation (2) is estimated with the restrictions $s_{NN} = s_{NZ}$ and $s_{ZN} = s_{ZZ}$. In this case, the geographically near pool coefficient is significant and larger than the geographically distant pool coefficient.

In order to check the robustness of the findings from the base-case, columns (2) through (4) report results from specifications that vary the definitions of geographic and

Table 4

Eighth-difference regression estimates for selected sensitivities of base case[a]

Variable	(1) Base case	(2) Geo. near ≤ 100 mi.	(3) Geo. near ≤ 200 mi.	(4) Tech. near = 3d SIC	(5) Corrected location	(6) Excl. hi-tech R&D
c	.080**	.077**	.082**	.068**	.081**	.115**
	(.031)	(.031)	(.031)	(.026)	(.031)	(.033)
l	1.06**	1.06**	1.06**	1.01**	1.06**	1.05**
	(.030)	(.030)	(.030)	(.024)	(.030)	(.031)
k	.029**	.029**	.030**	.040**	.029**	.063**
	(.015)	(.015)	(.015)	(.012)	(.015)	(.015)
$\sum k_{GnTn}$	.013**	.014**	.009**	.007**	.015**	.019**
	(.006)	(.005)	(.004)	(.002)	(.006)	(.009)
$\sum k_{GnTz}$	.005**	.002	.000	.000	.004*	.012**
	(.002)	(.002)	(.002)	(.002)	(.002)	(.004)
$\sum k_{GzTn}$	.014**	.015**	.015**	.008**	.014**	.010**
	(.002)	(.002)	(.002)	(.002)	(.002)	(.004)
$\sum k_{GzTz}$	-.000	.000	.000	-.001	.000	.006*
	(.002)	(.002)	(.002)	(.002)	(.002)	(.003)
adj R^2	.78	.78	.78	.85	.78	.76
N	1,484	1,484	1,484	1,484	1,484	1,484

[a] standard errors in parentheses; year-specific fixed effect coefficient estimates not reported.
* significant at the 10 percent level.
**significant at the 5 percent level.

technological nearness. Comparing column's (2) and (3) to the base case, the Σk_{GnTn} pool coefficient is biased down slightly as the definition of nearby firms is expanded to include those within 200 miles. Even at 200 miles, however, this coefficient is within one standard deviation of the base case point estimate. And the Σk_{GzTn} coefficient is unchanged as this pool is limited to firms beyond 200 miles. The Σk_{GnTz} coefficient, in contrast, becomes insignificant as the 'geographically near' boundary is expanded beyond 50 miles.

Comparing column (4) to (1), the technologically near pool coefficients fall as these pools are expanded to contain a firms entire three-digit SIC. In addition, the $\sum k_{GnTz}$ pool coefficient is insignificant as this pool is limited to only the most technologically distant firms in the panel.

Use of corporate headquarters to define firm location is one potential source of measurement error in the analysis. Presumably, a firm's R&D facility is the most likely location of spillover generation and reception. While R&D activity is often conducted at a headquarters location, this is not necessarily the case. The *Directory of American Research and Technology* was consulted to establish the reasonableness of the claim that corporate headquarters may be a useful proxy for the location of spillover generation and reception. This volume lists the location of a firm's corporate headquarters and R&D facilities. Approximately 60 percent of the firms included in this analysis have entries in at least one of the following volumes: 1975, 1977, 1983, 1988, 1993, 1998. Of this sample, 76 percent of firms conduct their R&D in the same city as their corporate headquarters. And a total of 87 percent conduct their R&D in the same state. In order to check the sensitivity of the base-case results to these variances, column (5) reports the results of the regression where the R&D facility location was used in place of the corporate headquarters location where these were found to differ. The results are unchanged from the base-case analysis.[21]

Finally, use of SIC 35 for this analysis may limit our ability to generalize these

[21] Of the 60 percent of firms observed in the directory, 24 percent conduct at least some of their R&D outside their headquarters city; 13 percent conduct R&D outside their headquarters state. Assuming these fractions apply to firms unobserved in the directory, an erroneous city is assumed for 50 firms, or less than 10 percent of the total panel; an erroneous state is assumed for 26 firms or five-percent of the total panel. These sample statistics are likely to represent an upper bound, however, if we assume that any underreporting in the directory is biased towards smaller firms that are more likely to conduct R&D at their headquarters location.

results if the relatively high-tech SIC 357, Computers and Office Equipment, are generating all of the spillovers. The results reported in column (6), which exclude SIC 357 R&D from spillover pools, suggest R&D spillovers are a general artifact of the panel. The most significant result of this sensitivity appears to be an upward bias in the own R&D coefficient, which is two standard deviations above the base case estimate. In addition, the Σk_{GnTz} coefficient is biased upward and nearly two standard deviations above the base case estimate, the Σk_{GzTz} coefficient is significant at the 10 percent level, and all the pool coefficients are estimated with lower precision.

4. CONCLUSION

This analysis improves our understanding of the importance of geographic and technological distance for identifying inter-firm R&D spillovers. Parameter estimates obtained in a production function framework indicate that spillovers are significant and important from geographically and technologically proximate R&D stocks. Evidence is presented which suggest an important measure of the apparent geographic localization of R&D spillovers may be a result of other factors that lead to industrial agglomeration.

The largest R&D spillovers are found to flow between firms in the same industry. However, spillovers within narrowly defined technological groups do not appear to be attenuated by distance. This is not to say that R&D spillovers necessarily play no role in agglomeration. While spillovers from outside a firm's narrowly defined industry are smaller, these spillovers do appear to be attenuated by distance. Consequently, R&D spillovers may play a role in the formation of industrially diverse agglomerations.

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.

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

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.

Cameron, Gavin, 1996, "Innovation and Economic Growth," *Centre for Economic Performance Discussion Papers* 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.

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.

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

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, 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.

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.

______________________________, 1999, "International Knowledge Flows: Evidence from Patent Citations," *Economics of Innovation and New Technology*, 8(1-2): 105-136.

Mairesse, Jacques and Bronwyn Hall, 1996, "Estimating the Productivity of Research and Development in French and United States Manufacturing: an exploration of simultaneity issues with GMM methods," in K. Wagner and B. van Ark, eds., *International Productivity Differences*, Elsevier.

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

Orlando, Michael J., 2000, "On the Importance of Geographic and Technological Proximity for R&D Spillovers: An Empirical Investigation," *CFederal Reserve Bank of Kansas City Research Working Papers* no. 00-02.

Pakes, Ariel and Zvi Griliches, 1984, "Patents and R&D at the Firm Level: a first look," in Zvi Griliches, ed., *R&D, Patents and Productivity*, Chicago: Chicago University Press.

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

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.

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.

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

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.