NBER WORKING PAPER SERIES

PUBLIC & PRIVATE SPILLOVERS, LOCATION AND THE
PRODUCTIVITY OF PHARMACEUTICAL RESEARCH

Jeffrey L. Furman
Margaret K. Kyle
Iain Cockburn
Rebecca M. Henderson

Public & Private Spillovers, Location and the Productivity of Pharmaceutical Research
Jeffrey L. Furman, Margaret K. Kyle, Iain Cockburn and Rebecca M. Henderson
NBER Working Paper No. 12509
September 2006
JEL No. L23, L65, O3, R3

ABSTRACT

While there is widespread agreement among economists and management scholars that knowledge spillovers exist and have important economic consequences, researchers know substantially less about the "micro mechanisms" of spillovers -- about the degree to which they are geographically localized, for example, or about the degree to which spillovers from public institutions are qualitatively different from those from privately owned firms (Jaffe, 1986; Krugman, 1991; Jaffe et al., 1993; Porter, 1990). In this paper we make use of the geographic distribution of the research activities of major global pharmaceutical firms to explore the extent to which knowledge spills over from proximate private and public institutions. Our data and empirical approach allow us to make advances on two dimensions. First, by focusing on spillovers in research productivity (as opposed to manufacturing productivity), we build closely on the theoretical literature on spillovers that suggests that knowledge externalities are likely to have the most immediate impact on the production of ideas (Romer, 1986; Aghion & Howitt, 1997). Second, our data allow us to distinguish spillovers from public research from spillovers from private, or competitively funded research, and to more deeply explore the role that institutions and geographic proximity play in driving knowledge spillovers.

Jeffrey L. Furman
Boston University – SMG
595 Commonwealth Avenue – #635a
Boston, MA 02215
and NBER
furman@bu.edu

Margaret Kyle
London Business School
Sussex Place
Regent's Park
London NW1 4SA
UNITED KINGDOM
and NBER
mkyle@london.edu

Iain M. Cockburn
NBER
1050 Massachusetts Ave
Cambridge, MA 02138
and NBER
cockburn@bu.edu

Rebecca Henderson
Sloan School of Management
MIT, E52-543
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
rhenders@mit.edu

I.   Introduction

Knowledge spillovers and their relation to the economics of research and development are central to much of the work of Zvi Griliches. In a sequence of papers, most notably "Issues in Assessing the Contribution of Research and Development to Productivity Growth" (Griliches, 1979) and "The Search for R&D Spillovers" (Griliches, 1992), Zvi outlined both the importance of understanding the nature and magnitude of spillovers for economic growth, and the formidable problems that complicate empirical work in the area. The importance of spillovers is well established. Their existence is fundamental to the theory of endogenous growth (Romer, 1986; Aghion and Howitt, 1997). Moreover, the assumption that publicly funded knowledge "spills over" from publicly funded institutions universities and laboratories is an integral part of the economic justification for their existence. The *measurement* of spillovers, however, is a complex and difficult undertaking.

Zvi identified three critical problems associated with measuring spillovers. The first is that of "distance." Zvi suggested that spillovers decline with the distance from the source to the potential recipient, and identified three sources of distance: market distance or relative position in the value chain; technological distance; and geographic distance. The second problem Zvi noted is that of lag structure, since there is little reason to believe that the spillover of knowledge is either instantaneous or uniform. Finally, the existence of spillovers affects the incentives of firms to do R&D. While spillovers *to* a firm are likely to make R&D investments more productive, and thus raise the incentives to invest in it, spillovers *from* the from the firm are likely to benefit competitors and thus lower the incentives to invest in research in the first place. Thus, spillovers and R&D are endogenous variables, a problem that makes identification significantly more challenging.

A stream of scholarly work, much of it by Zvi's students, collaborators and friends, built on this agenda to establish a consensus that, in Zvi's words "spillovers exist and are significant." For example, using a novel measure of technological distance, Jaffe (1986, 1989) found that firm-level R&D productivity was positively associated with the R&D investment of "technological neighbors" and that local (within state) university research increased rates of corporate patenting. Bernstein and Nadiri (1989) expanded on Jaffe's results by incorporating a more elaborate model of investment behavior, while Ward and Dranove (1995) and Adams (1990) experimented with the use of very long lag structures.

More recent work has attempted to expand Zvi's research agenda along two dimensions. The first is through the explicit recognition that the effectiveness of spillovers may be *mediated* — both by the types of investment made by the recipient firm (Cohen and Levinthal, 1989) and by the nature of the contractual relationships between the source and the potential recipient of knowledge. For example, Zucker, Darby and Armstrong (1998) argue that the positive correlation between the presence of "star scientists" in a location and the success of small biotechnology firms is driven not by simple geographic proximity, but rather by the existence of contractual relationships (specifically co-authoring) between "stars" and particular firms. Similarly, Cassiman and Veugelers (2002) find that if firms believe incoming spillovers are probably important, they are more likely to engage in cooperative R&D agreements.

The second line of work has attempted to link our understanding of spillovers to the geographic structure of multinational R&D. Research in international business has long been concerned with the tradeoffs that multinational firms face between being "home based" or "distributed" (Frost, 2001). In the context of the pharmaceutical industry, Chacar and

Lieberman (2003) and Furman (2003) argue that this tradeoff is particularly salient -- global pharmaceutical firms can organize their research efforts in a single site in order to maximize the advantages of internal scope and scale, or they can decentralize their research efforts with the aim of accessing knowledge spillovers that are only available across the globe.

In this paper, we explore the relationship between R&D productivity and the potential for spillovers provided by research local to distributed research laboratories, drawing on a sample of multinational pharmaceutical firms in the 1980s. To do this empirically, we match more than 8 million records of worldwide publications by public and private institutions at the level of the therapeutic class (market level) and employ a firm-therapeutic class-level panel database of drug discovery expenditures.

Our work differs from existing studies in a number of important respects. To our knowledge, with a single exception (Henderson and Cockburn, 1996), prior work in this area has analyzed spillovers at the level of the firm or of the laboratory, rather than at the level of the individual market, or product class. We also test for the presence of spillovers across the world, rather than in a single region such as the US or Europe, matching worldwide productivity to both locally and globally derived spillovers. Finally, we explicitly distinguish between publicly and privately generated spillovers. While several authors have explored the impact of privately generated knowledge on productivity (e.g., Jaffe, 1986; Cassiman and Veugelers, 2003), with the important exception of Adams (2002), efforts to compare the relative effect of the two are scant.

Consistent with expectations, we find that spillovers exist and are significant. Patent output at the therapeutic class level is positively correlated with a firm's "exposure" to papers related to that therapeutic class authored by scientists within thirty-five miles of

where the firm conducts research. Distinguishing between privately and publicly authored papers yields a striking result: patent output is positively and significantly correlated with publicly authored work, but negatively and significantly correlated with private sector work. Our results hold for both locally and globally generated work, with the effects being strongest for locally authored work but significant at both levels, and they are robust to alternative econometric specifications, as well as to the inclusion of controls for initial conditions in location and the geographic organization of the firm's research effort.

Our results suggest that, as many authors have suggested, public spillovers may play a major role in driving private sector productivity. This finding builds on Branstetter's (2003) evidence of significant spillovers from academic science to industrial patenting. We also obtain a result not evident in prior empirical work: private sector work may "crowd out" work by rivals or negatively affect their productivity, as some theories predict. This may be either because competitive success discourages rivalry, or because competitive success may literally "fish out the pool," making it harder and harder for rivals to discover new therapeutic entities or approaches. Taken together, our findings suggest that the problem of the optimal location of research and development for a multinational company is a multidimensional problem of considerable complexity, and that the relationship between spillovers and economic growth may be more nuanced than is sometimes assumed. We begin by briefly reviewing the relevant literature. Section III summarizes our data and gives some descriptive statistics. In Section IV, we outline our estimation approach and describe the results. Section V concludes and offers suggestions for future research in this area.

II. Literature Review

Knowledge spillovers and Marshallian externalities of all kinds are fundamental to theories in economic geography and to theories of endogenous growth. Marshallian externalities reduce production costs. They include the availability of specialized labor, the accumulation of human capital, the availability of specialized input services and other specialized infrastructure, and knowledge spillovers of all kinds, including face-to-face communication. The theoretical literature has generally not differentiated between the effects of privately and publicly generated spillovers, although Jacobs famously distinguished between localization effects flowing from the proximity of direct competitors and urbanization effects flowing from the volume of overall activity in the area (Jacobs 1969, 1984).

Factor mobility therefore plays a key role in all models of agglomeration externalities. As a broad generalization of a large and complex literature, with high factor mobility and/or easily transported products, the theoretical models predict strong clustering as a results of small differences in initial conditions. The endogenous growth models that incorporate location effects (see, for example, Aghion and Howitt, 1997 and Fujita and Thisse, 2003) combine core-periphery models of industrial location with Grossman-Helpman-Romer endogenous growth models with horizontally differentiated products. The R&D sector is modeled as using skilled labor to produce new varieties. In these models, externalities tend to be mutually reinforcing: more agglomeration leads to more productive R&D, which leads to more varieties, which leads to more demand, more concentration, more externalities and so on.

Broadly construed, the key result of these models is that the growth rate of the economy is driven by the spatial distribution of the innovation sector. With costless transfer of "patents" across regions, externalities and region-specific knowledge capital tend to pull all R&D labor into one location. Production may or may not follow depending on transportation costs. These theoretical models thus suggest that the existence and nature of spillovers have very significant implications for the nature of economic growth.

The finding that the spillover of knowledge is geographically localized is well-established (Autant-Bernard, 2003; Audretsch and Feldman, 1996; Feldman, 1994; Jaffe, 1989; Jaffe et al, 1993; Evenson and Kislev, 1973), as is the finding that university research has a positive effect on the productivity of local firms (Jaffe, 1989; Zucker and Darby, 1997). These results raise the question of the optimal location of large research efforts – since the choice to locate near a single cluster necessarily makes it more difficult to benefit from the research conducted at another. Large firms wishing to benefit from publicly funded research could respond by distributing their research efforts, but such a distribution is likely to be costly, since, as Chacar and Lieberman (2003) point out, a significant body of organizational theory predicts that there are very significant productivity benefits to the co-location of research. Allen (1977), for example, suggests that communication between researchers falls with the square of the distance between them, and although modern communications have made the creation of "virtual" teams feasible, managing them still creates formidable problems (Cummings, 2003a, 2003b).[1]

---

[1] Note that another key choice variable for the firm is whether to do research internally ("make" or integrate the research function) or rely on licensing the research efforts of others ("buy" research). However, all the firms we examine here performed a great deal of basic research internally, particularly during the time period we consider. We therefore do not deal with this choice explicitly, and take it as given for our purposes.

Moreover, work in industrial organization, particularly in the "racing" tradition, suggests that any single firm's choice of research location will also be shaped in complex ways by its rivals' choices. While the theoretical work in economic geography and in endogenous growth assumes that rivals generate positive externalities, industrial organization theory highlights the fact that increasing competition may trigger decreases in research investment or in competitive output (see Cockburn and Henderson, 1994, for a review of this complex and often contradictory literature and Aghion et al, 2003 for recent work exploring the relationship between market structure and investment in research).

The geographic organization of research in the global firm thus presents a unique setting in which to explore the consequences of geographically constrained spillovers for the organization of economic activity, and to explore the difference in roles (if any) between publicly and privately funded research in determining research productivity. The pharmaceutical industry is a particularly interesting context in which to explore this problem. As a number of researchers have noted, the industry is critically dependent on the outcomes of successful research to drive firm growth and profitability; consequently, it spends more on research, measured as a percentage of sales, than any other major industry (Gambardella, 1995). In addition, publicly funded research plays a major role in driving industry investment and productivity (Ward and Dranove, 1995; Cockburn and Henderson, 1998; Zucker and Darby, 1997). As a result, the tradeoff between the desire to locate all research in a single location and the drive to disperse it across the world to take advantage of local spillovers is particularly acute for multinational pharmaceutical firms.

The implications of this tradeoff for the geographical boundaries of the firm are of particular interest for emerging research on the strategic management of multinational

corporations.[2] Chacar and Lieberman explore this tradeoff using data aggregated to the firm level from 21 US based pharmaceutical companies. They regress the number of NCEs (New Chemical Entities, a measure of research output) and patents against a number of measures of geographic structure, including whether the firm operates only a single lab and the number of foreign labs operated by the firm, and find that firms with more foreign labs appear to be more productive than those with fewer. They interpret this result as suggesting that locally captured spillovers may play an important role in driving firm productivity. Similarly, Penner-Hahn and Shaver (2005) find that Japanese pharmaceutical firms with international R&D facilities produce a larger number of patents than firms with purely domestic research facilities.

We expand on these approaches in a number of ways. We analyze the productivity of a firm's research effort at the level of the therapeutic class, rather than looking at the total output of all the firm's research efforts. This level of analysis permits more precise comparisons between firms' research programs as well as within-firm productivity differences across therapeutic areas that may be driven by local spillovers. In addition, we test for the importance of local spillovers by constructing local measures of knowledge generation directly, rather than by inferring their impact from differences in the geographic construction of R&D. Importantly, we explicitly distinguish between privately and publicly generated spillovers, allowing us to address industrial organization theories on competitive racing in R&D as well as the endogenous growth and economic geography literatures.

---

[2] It is useful to note that additional work on geographic location in international business research examines qualitative issues associated with organizing and managing distributed firms (e.g., Ghoshal, 1993) location-specific differences in the organizing strategy of R&D facilities (Furman, 2003), and the determinants of R&D directed FDI decisions (e.g., Chung and Alcacer, 2004); existing work does not, however, identify a link between distributed R&D location decisions and productivity.

III. Database Construction

To accurately measure drug discovery productivity, we require measures of outputs and inputs at the level of the therapeutic class–year. If it were possible to obtain ideal data, we would want to conduct this analysis at the laboratory–therapeutic class–year level. Although it is possible to attribute outputs to individual laboratories and to measure knowledge-based inputs in the region local to specific laboratories, accurate measures of drug discovery expenditures are not available at the level of the individual facility. As a consequence, we analyze drug discovery productivity at the firm-therapeutic class-year level; we do, however, incorporate facility-specific data on knowledge inputs into the production function we estimate.

We measure drug discovery output with PATENTS, a count of the number of "important patents" granted in each firm - therapeutic class - year.[3] Patents are considered important if they have been granted in any two of the three major markets in the world (the European Union, Japan, or the U.S.). We assign patents to the year of application. In focusing on patents-based measures of drug discovery output we build squarely on prior work in this area (Gambardella, 1995; Henderson and Cockburn, 1996; Penner-Hahn and Shaver, 2004). We recognize that patents are one of many potential measures of outputs from the drug discovery process, including new chemical entities (NCEs), investigational new drug applications (INDs), and new products introduced and note that prior work has demonstrated a significant correlation between PATENTS and INDs and other measures of drug discovery output.

---

[3] We rely on Derwent's categorization of patents into therapeutic classes. See Henderson and Cockburn (1996), p.57, for additional information on patent data from Derwent.

Part of the novelty of our paper derives from the level of detail with which we measure inputs into the drug discovery process. Our methodology reflects both the extent of capital devoted to drug discovery research at the therapeutic class level as well as the extent of knowledge inputs external to the firm but present in the regions geographically proximate to its research facility. Our measure of drug discovery research spending, DISC_IJ, represents millions of constant 1986 dollars of expenditure on drug discovery efforts in each firm - therapeutic class - year. This variable is designed in order to measure the amount of expenditures devoted to identifying compounds with promising medicinal effects. It excludes expenditures directed at development or other R&D activities. The construction of DISC_IJ and its classification into therapeutic classes follows the descriptions of Henderson and Cockburn (1994 and 1996). For each therapeutic class, we consider both a stock and a flow of DISC_IJ. Stocks are computed based on the assumption that knowledge depreciates at an annual rate of 15% (Griliches, 1994). We include the flow of expenditures to capture *current* differences in the research efforts of firms that are not picked up by differences in stocks. For example, we might expect the research productivity of a firm that is beginning to ramp up a research program to be higher than that of a firm that is gradually ending a program 20 years old.

To capture knowledge inputs into the drug discovery process at the level of the therapeutic class, we develop novel and detailed bibliometric measures of the extent and distribution of publicly-available knowledge in the life sciences. We draw on the National Library of Medicine's PUBMED database and the Institute for Scientific Information's Science Citation Index (SCI) to construct these instruments. The National Library of Medicine's PUBMED database includes bibliographic information from more than 4,600

journals in fields of biomedicine and the life sciences. The database covers journals from more than seventy countries. It reports detailed information on each publication, including paper titles, author names and affiliations and information on the citations made and citations received.

To classify publications by therapeutic class, we exploit another feature of the data from the PUBMED database. This database classifies each publication using terms (similar to Journal of Economic Literature terms) that characterize the themes and content of each article. These descriptors, called Medical Subject Headings (MeSH), are organized hierarchically and categorically, and each paper is associated with multiple MeSH categories. For example, a paper that examines the effect of interfering with the renin angiotensin cascade in rats will be associated with MeSH headings for research on (a) rats, (b) hyptertension, and (c) the renin angiotensin cascade, as well as a number of other headings that characterize other elements of the research design and physiological and anatomical systems being studied. We focus on those MeSH headings that refer to the therapeutic classes of research addressed by particular papers.[4] Thus, using the PUBMED database allows us to create a useful indicator of the extent of research conducted in various therapeutic classes.

The PUBMED data do not, however, provide extensive information about the *location* in which the research is done. Until recently, PUBMED has only listed the affiliation and address for each paper's first author. In many cases, the leading researcher on life sciences papers is listed as the last author on the paper rather than the lead author. As a result, we undertake a significant data effort to generate more complete information on the

---

[4] These include branches "C" and "D" of the MeSH hierarchy; details can be found on the PUBMED website at http://www.nlm.nih.gov/mesh/meshhome.html.

geographic location of life sciences research. In particular, we draw on the set of publications chronicled by the Science Citation Index (SCI). The SCI catalogues publications in nearly 5,000 international academic and industry journals, identifying, among other relevant characteristics, authors' names, addresses, and institutional affiliations. Moreover, SCI data list each unique address and institutional affiliation associated with each paper. Unfortunately, addresses are not matched to authors and each address appears only once per paper, even if more than one of the paper's authors are associated with that address. In this paper, we consider each address as a separate observation. Note that by using the address field in addition to the affiliation, we are able to match papers to the specific R&D laboratory within a firm, not merely to the firm's headquarters.

In order to take advantage of the address affiliations (i.e., geographic information) in the SCI and the therapeutic class information available in PUBMED, we create a dataset that consists of the set of papers that appear in both sources.[5] Because the overlap between the sets of journals is not perfect, we lose a significant number of journals in the matching process. However, we believe that a dataset that consists of only journals in both datasets has conceptual as well as practical advantages. The Science Citation Index includes a number of scientific journals outside the life sciences, while PUBMED contains a number of journals in the medical field that are not scientific.[6] The intersection of the two databases is, therefore, more precise in identifying journals whose articles may be considered as useful knowledge inputs into the drug discovery research process.

We begin with the list of journals that appear in both the SCI and PUBMED datasets. Using this list, we identify 4.0 million papers in the SCI that could also appear in PUBMED

---

[5] We merge these based on journal names and paper titles (where both are purged of all non-text characters).
[6] An example is a journal called *Ontario Nurse*, which is unlikely to publish cutting-edge research in biology.

between 1981 and 1990. Matching based on journal name, paper title, and authors' names, we are able to identify 2.3 million papers that appear in both data sources. We lose potential matches (a) because journals do not always appear in both data sources at the same time and (b) because, despite our best efforts, differences in spelling and word use persist across data sources. For example, SCI and PUBMED differ on occasion in their use of British vs. American English, their use of text numerals vs. numbers, and their use of Greek and other foreign language characters.

In order to classify articles by therapeutic class, we draw on the MeSH classification scheme maintained by the NLM. Approximately 35 percent of all articles fall into MeSH headings that suggest relevance to a particular therapeutic class of drug discovery research. Articles that are not classified into therapeutic class include those written in for a more broad or basic scientific audience, as well as those that are not directly related to drug discovery research.

Using the address information available from the SCI, we also classify each article-address in our dataset geographically. Article-addresses associated with firms in our dataset are assigned to the particular firm-laboratory from which they originated. Thus, we know the geographic distribution of articles within each firm as well as across firms. We treat each address as a "knowledge observation" and are able to identify approximately 2.9 million paper-addresses in our data. In cases with a large number of co-authorships within the same organization (or address), this method will undercount the number of author-publications because the SCI address field includes only 'unique addresses,' (i.e., it lists all of the unique address affiliations of the authors, but does not list the same address-affiliation multiple times if there are multiple authors from that address). If the tendency to co-author

with "insiders" differs substantially and systematically across regions, our measure could introduce a bias. For example, our measure would assign more "knowledge observations" to a region with a total of 50 biologists split between two organizations than to another region with a total of 50 biologists all at the same institution. However, we do not have prior information that suggests that such a bias is likely to affect our results in a serious way.

The SCI address information enables us to assign each article-address to a "life sciences region." Because the definitions for administrative boundaries (such as states, provinces, MSAs, or EU-NUTS regions) differ greatly across countries, we create regions whose definitions are based on the locations that generate the greatest number of life sciences publications. Specifically, each city that produces more than 5,000 life sciences articles (equivalent to approximately 0.15% of the world's total) is considered to be a regional "centroid." For each region, we count the number of articles published by authors whose affiliations indicate an address within the 35-mile radius around the regional centroids.[7] (We accomplish this by using databases provided by the U.S. National Imagery and Mapping Agency that list latitudes and longitudes for world cities and towns.) Where high publication cities are geographically proximate, we define the region based on the city with the larger number of publications – e.g., Leverkusen and Cologne, Germany both produce more than 5,000 publications, but since Leverkusen produces more publications, it is considered as the center of its region with Cologne as one of its local cities.

To further categorize paper-addresses as either "public" or "private" science, we use the affiliation field. All academic, government, and non-profit affiliations are considered "public" science. These account for 90% of the total. Of these, approximately two-third of

---

[7] Increasing the radius to 50 and 100 miles had no major effect on the results we report here.

public articles come from universities, while the remaining one-third come from medical schools and government sources. We define "private" science as all other affiliations.

IV. Data Description

The names, definitions, and descriptive statistics for our key variables appear in Table 1. We draw these data for thirteen therapeutic classes from nine of the largest ethical pharmaceutical firms between 1981 and 1990. As not every firm conducts research in all therapeutic classes in each year, we have an unbalanced panel totaling 704 observations.[8] Our principal measure of drug discovery output is PATENTS. On average, a firm is granted nearly three PATENTS per year in the therapeutic classes in which it is active. This measure is quite skewed, with many classes receiving zero patents in any one year. Driving such patenting, the average firm expenditure per therapeutic class–year is $4.1 million (1986 dollars). (Note that this is significantly greater than the average level reported in Henderson and Cockburn (1996) because we define therapeutic class at a higher level of aggregation.) Depreciating expenditures and patents at a rate of 15% per year, the average stock of PATENTS is 16.2 and the average stock of discovery expenditures (LAG DISC_IJ_S) is $11.7 million. Across all therapeutic classes, the firms in our sample spent $47 million on discovery research annually and published 38.2 papers per therapeutic class per year.[9]

---

[8] As has been noted in prior work using these data, the research program – rather than the therapeutic class – is the ideal level at which to measure drug discovery productivity (Henderson and Cockburn (1996)). While our data classification scheme is sufficiently sophisticated to classify publications data at the therapeutic class level, we cannot data at the level of the research program. The therapeutic class codification scheme is sufficiently broad that each class may include multiple research programs that are not directly related to one another. Combining research programs in this way adds noise to the measurement of inputs and outputs at the therapeutic class level; we do not believe, however, that it induces biased productivity estimates.

[9] Because data on publications are not available prior to the start of the sample period (in 1981), we compute stocks as they accumulate in the sample. To correct for initial conditions in the local knowledge stock, we examine specifications that include a variable that reflects the initial count of public and private publications in each region.

Our sample firms differ markedly in the extent to which they invest in each particular therapeutic class, as well as in the number of publications and patents they receive (Table 2A). While some firms maintain only one laboratory, many firms in the sample maintain laboratories in multiple regions (and countries). For these firms, there are substantial differences across laboratories in extent to which they perform research in various therapeutic classes (Table 2B).

Likewise, there is considerable variation across countries and regions within country in the extent to which they publish life sciences research. The majority of research emanates from North America, Western Europe, Japan, and Australia. While this correlates with the extent of industrial development, it is striking in light of the distribution of world population, which has much greater concentration in South America, Africa, India, and Southeast Asia. The distribution of research by therapeutic class varies across location. For example, Cardiovascular research is concentrated in the United States and Western Europe, while Parisitology research is performed in a wider set of countries, including many in tropical areas. (Figures 1 and 2 depict the world distribution of Cardiovascular and Parisitology research, respectively.)

Geographic concentration in research is also evident within-country in our sample. For example, life sciences research in the United States is most extensive in California, Massachusetts, and in the area between Washington, DC and Baltimore. It is also important to note that the the geographic distribution of PRIVATE life sciences research differs somewhat from that of PUBLIC research. For example, relatively more PRIVATE research is performed in New Jersey and Washington during the sample period, while relatively more PUBLIC research takes place in Ohio, North Carolina, and Georgia.

To evaluate the influence of external research on internal productivity, we compute EXPOSURE, which measures the number of publications within a 35-mile radius of each laboratory in which the firm conducts its research. This variable is designed to capture the magnitude of potential spillovers from life scientists outside the focal firm in each period. Firms can increase their EXPOSURE to available research by locating in areas in which there are many other private or public institutions engaged in life sciences. By measuring EXPOSURE at the firm-therapeutic class level, we account for a particular location's proximity to valuable science in one research area, but distant from science in other areas. On average, each therapeutic class is proximate to 33,000 papers authored by individuals affiliated with PUBLIC institutions (universities, government agencies, or other not-for-profit entities) and fewer than 1000 papers authored by individuals at PRIVATE sector firms.

A second element of closeness to external knowledge is whether a firm's geographic allocation of research in each therapeutic class matches the external geographic distribution of research in that class. Developing measures that adequately compare the geographic distribution of firm–therapeutic class research with the geographic distribution of therapeutic class-level research outside the firm is extremely complex methodologically. As a preliminary way of exploring this issue, we experimented with a straightforward measure of the match between the firm and worldwide distribution of class-specific research, although we did not obtain conclusive results.[10]

---

[10] We experimented with the variable GINI, which we computed as the mean squared difference over locations between firm i's share of its total activity in each location and the world share of activity in each location (Krugman, 1991; Audretsch and Feldman, 1996). More formally, $GINI_{ijt} = \Sigma(s_{ijt}-x_{jt})^2$, where X_{jt} is the 1 x K vector of shares of each location k in the world total of activity in therapeutic class j and S_{ijt} is the 1 x K vector of the share of each location k in firm i's total activity in therapeutic class j (each in time t). As smaller GINI values correspond to a closer match between the firm's geographic allocation of science, we would expect a negative coefficient on GINI in the productivity regressions.

V. Empirical Results

We draw on the "production function" approach to the measurement of productivity, in which patents are generated as a function of the research investment of the firm and of its knowledge capital, in the manner of Griliches (1979, 1994, 1995). We estimate a function of the form:

$$\text{Log } Y_{i,j,t} = \alpha(\log X_{i,j,t}) + \beta(\log K_{i,j,t}) + A_{i,j,t}(t) + u_{i,j,t} \tag{1}$$

where Y is a measure of output at the firm (here, patents at the therapeutic class level), X is a vector of inputs (here, current research spending), and K is a measure of cumulated research effort and i, j, and t, index firm, therapeutic class, and year, respectively. A(t) represents other factors which affect output and change systematically over time. Included in this is, for example, exposure to local scientific and technical knowledge and measures of firm organization, such as the geographic dispersion of the firm. Lastly, u reflects all other random unsystematic fluctuations in output.[11]

This is a reduced-form specification, and thus the interpretation of parameter estimates is subject to many caveats. For example, a firm's patenting behavior may well be influenced by the R&D spending of its rivals. Certainly, the selection of therapeutic classes in which a firm focuses research efforts is a source of endogeneity, and one we make no pretense of correcting for. Here, we think the greater concern in the context of our research question is the endogeneity of location choice. A firm's choice of research location, and therefore exposure to local spillovers, could also be a function of its desire to avoid spilling over to competitors. While a firm's research in a locality may affect the efforts of its competitors in the same geographic area, we think it is reasonable to assume that public

---

[11] The functional form of this equation should be interpreted as a first approximation to a much more complex relationship. Most studies that use this technique explore the degree to which the results are sensitive to functional form. They are usually not.

research in a locality is exogenous. Most universities are older than the labs in our sample, and the location of government laboratories was also fixed during this period. If a very productive firm began opening labs in additional locations (or assumed control of addition labs through a merger), then its exposure to local science would increase and we might erroneously attribute high productivity to spillovers from local science. During our sample period, locations are largely fixed: firms did not open new labs or close existing labs. Only one large merger occurred between firms in our sample, close to the end of the sample period, which we believe would have only a small effect.

Another potential problem we have in identifying spillovers is that both publications and patenting in a therapeutic class could be affected by exogenous shocks to technological opportunities for that class. This affects our interpretation only if research in therapeutic classes is geographically concentrated, as our discussion above indicates. To address this, we include therapeutic class fixed effects in one specification, which essentially control for the initial geographic distribution of research in a class. The coefficients on the other parameters of interest then reflect within-class productivity differences across locations, netting out any technological shock to the class. However, this does not control for correlated temporary shocks, which may still be of concern for the analysis and interpretation of the results.

Tables (3) and (4) present our results. Since our dependent variable is often truncated at zero, in Table (3) we model equation (1) using a Poisson specification. As this has the well-known problem that it imposes the condition that the mean of the distribution is equal to the variance, we replicate the analyses in Table (4) using Negative Binomial regressions. Our key results are robust to either specification. We include year fixed effects

in every specification to control for fluctuations in overall rates of patenting and for any general common trend in "propensity to patent" across the industry.

Model (1) presents the "base case," including only our flow and stock measure of research (discovery) spending. The coefficients are positive and highly significant, with magnitudes reassuring close to those that have been observed in prior work. The differences in the magnitudes of the coefficients note that historical investments in research programs have continued effects on the productivity of current efforts.

Model (2) includes our measure of local spillovers, EXPOSURE_FLOW. Recall that this variable is defined as the total number of papers (from both public and private institutions) published within a 35-mile radius of the firm's research activity in a particular therapeutic area. Consistent with prior research, this coefficient is positive and significant, suggesting a positive association between the extent of local knowledge and research productivity.[12]

Model (3) breaks this measure of local spillovers in two, as a function of the institutional affiliation of the author: EXPOSURE_FLOW_PUB and EXPOSURE_FLOW_PRIV. EXPOSURE_FLOW_PUB retains its positive and significant coefficient, but EXPOSURE_FLOW_PRIV is negative and significant. This is our core result, and holds across all models and both econometric specifications.

This is a striking result. The positive impact of academic research on productivity is expected, and entirely (and reassuringly) consistent with prior research. The negative effect

---

[12] In additional, unreported models, we also experimented with including GINI, our measure of the degree to which the firm's distribution of research activity in a particular class mirrors that of the general distribution, as an additional indicator of a firm's closeness to life science research. GINI was not, statistically significantly in any of these models. This may suggest that matching the world wide research distribution is not correlated with research productivity; however, this result may also arise because this measure is not sufficient in capturing the phenomenon of interest.

of privately generated knowledge on productivity is more surprising. Adams (2002) found that laboratory productivity was correlated with both publicly and privately generated knowledge, and the majority of the theoretical literature assumes that spillovers across private firms have positive effects. However, our result is consistent with Jaffe (1986)'s analysis that controlled for technological but not for geographical distance, which found that private sector research spending was positively correlated across firms, but profits were negatively correlated with rival's R&D spending.

This is a reduced form result, so it probably summarizes a complex mix of factors. It may be that extensive rival publication in a particular area reduces the opportunity for any single firm to make the novel discoveries that are a prerequisite to patent generation. Alternatively, extensive rival publication may signal significant future competition in a particular market, thus reducing the attractiveness of effort in the area and the incentive to generate patents after research efforts have already been expended. Indeed, there are likely numerous credible explanations for this finding and we believe that clarifying this result is a useful goal for future research.

Model (4) introduces N_CENTROIDS, or the number of centroids in which the firm (in this particular therapeutic class) has located research activity, another measure of the degree to which the firm is attempting to take advantage of research across the world. It is significant and negative in the Poisson results, although insignificant in the negative binomial specification. (N_COUNTRIES, another measure of geographic dispersion with which we experimented, is also negative and significant in both specifications when included in the model). We interpret this result as consistent with the hypothesis that firms

that are "stretched too thin" fail to take advantage of local economies of scope and suffer a reduction in productivity in consequence.

In model (5), GLOBAL_PUB_STOCK, the stock of *globally* generated academic papers is positive and significant in both specifications, while GLOBAL_PRIV_STOCK, the stock of globally generated private papers is negative and significant in both specifications. This result is consistent with our local result, and suggests that firms benefit (or suffer) from global spillovers as well as from local. The results from this model suggest that the marginal impact of an additional 1000 local publications is 0.19 PATENTS, while the addition of 1000 publications to the global stock of relevant science yields .11 additional PATENTS. Thus, the implied elasticities associated with these measures are consistent with research in economic geography and endogenous growth that suggest that local knowledge is more likely to spill over than distant knowledge.

Model (6) introduces LOCAL_START_PUB and LOCAL_START_PRIV. These controls attempt to address the endogeneity of research location by controlling for "initial conditions" in each location. Their signs are consistent with our flow results, which are robust to their inclusion. It is important to note, however, that the inclusion of these variables does not control for the possibility of temporary shocks that may also affect location decisions. Model (7) presents the full model and model (8) introduces therapeutic class fixed effects. Introducing therapeutic class fixed effects alters the interpretation, sign, and significance of the initial conditions variables, although it does not affect the paper's core results. Models without therapeutic class fixed effects exploits differences across locations in public science. Much of those differences are driven by specialization in therapeutic areas. Removing those differences in specialization by including therapeutic

class FEs appears to leave too little variation over time or across locations to get a statistically significant (or economically meaningful) estimate, except in the case of LOCAL_START_PUB in the POISSON model, which is negative and statistically significant. This unexpected finding may suggest that patenting productivity was highest in those therapeutic classes about which the least was known at the beginning of the period.

In addition to the models presented here, we also experimented with a number of alternative specifications, including those that explored the effects of clustering errors by therapeutic class and by firm, and with random and firm fixed effects models. While our results are robust to error clustering (the standard errors increase, but most coefficient remain significant at the 5% level), they are not robust to the use of random or firm fixed effects. Given the very limited extent to which the geographic distribution of research changes within the firm during the period covered by our study, this is not, perhaps, a surprising result. In this period, the geographic location of research is the firm fixed effect.

VI. Discussion

To our knowledge, ours is the first paper to explore the impact of locally generated spillovers across the entire world on the research productivity of the multinational firm, and among a small set of papers that distinguishes explicitly between the effects of publicly and privately generated knowledge. We find that locally generated knowledge is strongly correlated with research productivity, but that this positive correlation is driven overwhelmingly by the impact of publicly generated knowledge. Privately generated knowledge appears to be negatively correlated with productivity, a result consistent both

with a "crowding out" hypothesis and with the hypothesis that the anticipation of future rivalry reduces the incentive to patent in any given area.

Our results have intriguing implications both for our understanding of economic geography and for our understanding of the global organization of research in the multinational firm. They suggest that there may be a complex tradeoff between the desire to locate close to sources of academic knowledge on the one hand and the desire to avoid competing too closely with rivals on the other and that it is possible that over dispersing research activities may have a negative effect on productivity. They also imply that models that fail to distinguish between the nature of private and public spillovers may be misspecified.

Further work in this area could move in a number of directions. One is the attempt to model the endogeneity of research location more explicitly, perhaps through the use of suitable instruments and the more sophisticated structural modeling of investment choices. Such an effort will be a challenging undertaking, given the difficulty of finding appropriate instruments and the fragility and complexity of theoretical models of the problem. The use of data from more recent periods that feature much greater variation in the geographic location of research may help to address these issues. It is also surprising that, if local spillovers from public science are so strong, pharmaceutical firms had chosen to locate in New Jersey, for example, to a greater extent than Massachusetts or the Washington, DC area. However, changes since the end of our sample period in the 1990s and early 2000s, such as the growth of the biotech industry around Boston and the decision by Novartis to locate its research headquarters in Cambridge, suggest that firms may have, indeed, begun to respond to opportunity to appropriate local spillovers from public institutions.

Another possibility is to explore whether local spillovers are mediated in their effects, as Zucker and Darby and their collaborators and Cassiman and Veugelers suggest. Public research in the medical sciences is very widely disseminated, and in some respects it is surprising to find that it has a strong local impact. Is the positive effect of local research a function of institutional ties between firms and local institutions? Using our data to identify co-authoring relationships between institutions may shed some light on this question, and add to the recent findings of Singh (2003) and Breschi and Lissoni (2003). Similarly, the negative effect of competitive work, if it exists, should have global impact – why does the local effect appear to be so strong? Additional work could explore the micro-mechanisms that lie behind this result, further illuminating our understanding of the dynamics of the geographic organization of research and the relationship between spillovers and economic growth.

References

Adams, J. 2002. Comparative Localization of Academic and Industrial Spillovers. *The Journal of Economic Geography* 2: 253-278.

Aghion, P. and P. Howitt. 1997. *Endogenous Growth Theory*. Cambridge: MIT Press.

Aghion, P., R. Blundell, R. Griffith, and P. Howitt. 2004. Entry, Innovation and Growth: Theory and Evidence, working paper.

Allen, T.J. 1977. *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R&D Organization,* Cambridge: MIT Press.

Audretsch, D. and M. Feldman. 1996. Knowledge spillovers and the geography of innovation and production. *American Economic Review*, 86(3): 630-640.

Autant-Bernard, C. 2003. Specialisation, Diversity and Geographical Diffusion of Knowledge, mimeo, CREUSET – Université Jean Monnet.

Bernstein, J.I., and M.I. Nadiri. 1989. Research and Development and Intra-industry spillovers: An empirical application of dynamic duality. *Review of Economic Studies*: 249-69.

Branstetter, L. 2003. Measuring the Impact of Academic Science on Industrial Innovation: The Case of California's Research Universities, mimeo, Columbia University.

Breschi, S. and F. Lissoni. 2003. Mobility and Social Networks: Localised Knowledge Spillovers Revisited, CESPRI Working Paper, Universita Bocconi, Milan.

Cassiman, B. and R. Veugelers. 2002. R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. *American Economic Review* September 2002: 1169-1184.

Chacar, A. and M. Lieberman. 2003. Organizing for Technological Innovation in the U.S. Pharmaceutical Industry. *Advances in Strategic Management,* J.A.C. Baum and O. Sorenson. JAI/Elsevier: Oxford, UK, 20: 317-340.

Chung, W. and J. Alcacer 2004. Knowledge Seeking and Location Choice of Foreign Direct Investment in the United States. *Management Science* 48(12): 1534-1554.

Cockburn, I. and R. Henderson. 1994. Racing to Invest? The Dynamics of Competition in Ethical Drug Discovery. *Journal of Economics and Business Strategy* 3(3): 481-519.

Cockburn, I. and R. Henderson. 1998. Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery. *Journal of Industrial Economics* 66(2): 157-182.

Cohen, W.M. and D.A Levinthal. 1989. Innovation and Learning: The Two Faces of R&D. *Economic Journal*. 99: 569-596.

Cummings, J. 2003. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science.* 50(3): 352-364.

Cummings, J. 2003. Leading groups from a distance: How to mitigate the consequences of geographic dispersion. In *Leadership at a distance* ed. S. Weisband & L. Atwater.

Evenson, R.E and Y. Kislev. 1973. Research and Productivity in Wheat and Maize. *Journal of Political Economy*. 81(6): 1309-1329.

Feldman, M. 1994. *The Geography of Innovation*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Feldman, M. and Y. Schreuder. 1996. Initial Advantage: the Origins of Geographic Concentration of the Pharmaceutical Industry in the Mid-Atlantic Region. *Industrial and Corporate Change,* 5: 839-862.

Frost, T.S. 2001. The Geographic Sources of Foreign Subsidiaries' Innovations. *Strategic Management Journal*. 22: 101-123

Fujita, M. and J.F. Thisse. 2003. Does Geographical Agglomeration Foster Economic Growth? And Who Gains and Loses from It? *Japanese-Economic-Review*. 54(2): 121-45.

Furman, J.L. 2003. Location and Organizing Strategy, *Advances in Strategic Management*, ed., J.A.C. Baum and O. Sorenson. JAI/Elsevier: Oxford, UK, 20: 49-88.

Gambardella, A. 1995. *Science and Innovation in the US Pharmaceutical Industry.* Cambridge: Cambridge University Press.

Ghoshal, S. 1993. *Organizational Theory and Multinational Corporations*. London, UK: Macmillan.

Griliches, Z. 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*. 10 (1): 92-116.

Griliches, Z. 1992. The Search for R&D Spillovers. *The Scandinavian Journal of Economics*, 94(Supplement): 29-47.

Griliches, Z. 1994. Productivity, R&D and the Data Constraint. *American Economic Review*. 84(1): 1-23.

Griliches, Z. 1995. R&D and Productivity: Econometric Results and Measurement Issues, Chapter 3 in *Handbook of the Economics of Innovation and Technical Change*, Ed. P. Stoneman, Ed. Oxford, UK: Blackwell.

Guillain, R. and J.M. Huriot. 2001. The Local Dimension of Information Spillovers: A Critical Review of Empirical Evidence in the Case of Innovation. *Canadian Journal of Regional Science*. 24(2): 313-37

Henderson, R. and I. Cockburn. 1994. Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal*. 15: 63-94.

Henderson, R. and I. Cockburn. 1996. Scale, Scope and Spillovers: The Determinants of Research Productivity in Drug Discovery. *Rand Journal of Economics*. 27(1): 32-59.

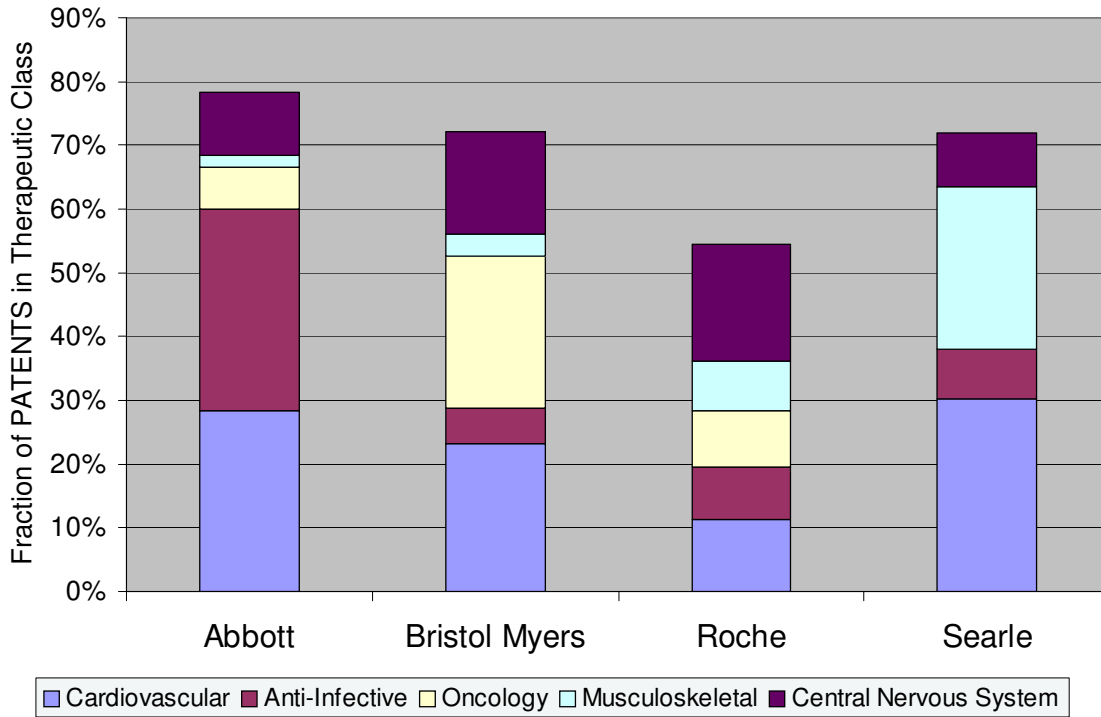Jacobs, J. 1969. *The Economy of Cities*. London, UK: Penguin.

Jacobs, J. 1984 . *Cities and the Wealth of Nations: Principles of Economic Life*. New York: Vintage.

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

Jaffe, A.B. 1989. Real Effects of Academic Research. *American Economic Review*. 79: 957-970.

Jaffe, A.B., M. Trajtenberg, and R.M. Henderson. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*. 434: 578-598.

Krugman, P. 1991. Increasing returns and economic geography. *Journal of Political Economy*. 99(3): 483-499.

Monjon, S. and P. Waelbroeck. 2003. Assessing spillovers from universities to firms: Evidence from French firm-level data. *International Journal of Industrial Organization.* 21(9): 1255-1270.

Penner-Hahn, J. and J.M. Shaver. 2005. Does international research and development affect patent output? An analysis of Japanese pharmaceutical firms. *Strategic Management Journal.* 26(2): 121-141.

Porter, M.E. 1990. *The Competitive Advantage of Nations*. New York: The Free Press

Romer, P. 1986. Increasing Returns and Long-Run Growth. *Journal of Political Economy*. 94(5): 1002-1037.

Siegel, D.P. Westhead and M. Wright. 2003. Asessing the impact of university science parks on research productivity: Exploratory firm level evidence from the United Kingdom. *International Journal of Industrial Organization*: 21(9): 1357-1369

Singh, J. 2003. Social Networks as Drivers of Knowledge Diffusion. mimeo, INSEAD-Singapore.

Spence, M. 1984. Cost Reduction, Competition and Industry Performance. *Econometrica*. 52(1): 101-121.

Ward, M. and D. Dranove. 1995. The Vertical Chain of R&D in the Pharmaceutical Industry. *Economic Inquiry*. 33: 1-18.

Zucker, L.G., M.R. Darby and M. Brewer. 1998. Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises. *American Economic Review*. 88(1): 290-306.

Zucker, L.G, M.R Darby, and J. Armstrong. 1998. Geographically Localized Knowledge: Spillovers or Markets? *Economic Inquiry*. 36: 65-86.

Table 1
Variable Definitions and Descriptive Statistics

Variable	Definition	Mean	Std. Dev.	Min.	Max.
Firm-Class-Year variables (*representing data for firm* i, *in therapeutic class* j, *in time* t)					
PATENTS	"Important" patents	2.96	4.56	0.00	34.00
KPATS	Beginning of year stock of PATENTS*	16.22	18.66	0.00	127.21
DISC_IJ	DISCOVERY FLOW 1986$M	4.14	5.66	0.00	48.67
DISC_IJ_S	Beginning of year stock of DISC_IJ*	11.70	15.56	0.00	105.51
LOG DISC_IJ	Log of DISCOVERY FLOW 1986$M	0.56	1.78	-8.16	3.89
LOG DISC_IJ_S	Log of Beginning of year stock of DISCOVERY*	1.15	2.48	-12.60	4.66
KPUBS_OWN	Stock of own papers in class*	38.24	53.41	0.00	373.59
N_CENTROIDS	Number of centroids in which a firm laboratory doing research in t-class *j* is located	1.90	1.28	0.00	8.00
EXPOSURE_FLOW_PRIV	Flow of PRIVATE papers a firm is exposed to in the areas around its labs	0.07	0.20	0.00	2.23
EXPOSURE_FLOW_PUB	Flow of PUBLIC papers a firm is exposed to in the areas around its labs	3.16	3.26	0.00	18.67
EXPOSURE_STOCK_PRIV	Stock of PRIVATE papers a firm is exposed to in the areas around its labs*	1.29	1.47	0.00	9.74
EXPOSURE_STOCK_PUB	Stock of PUBLIC papers a firm is exposed to in the areas around its labs*	9.32	10.06	0.00	56.87
LOCAL_START_PRIV	Stock of PRIVATE papers a firm is exposed to in the areas around its labs at start of period*	0.41	0.55	0.00	2.72
LOCAL_START_PUB	Stock of PUBLIC papers a firm is exposed to in the areas around its labs at start of period*	2.86	3.50	0.00	16.62
GLOBAL_PRIV_FLOW	Flow of PRIVATE papers in a therapeutic class	1.82	1.62	0.24	9.60
GLOBAL_PRIV_STOCK	Stock of PRIVATE papers in a therapeutic class*	4.96	4.64	0.53	31.76
GLOBAL_PUB_FLOW	Flow of PUBLIC papers in a therapeutic class	16.36	11.95	1.08	69.30
GLOBAL_PUB_STOCK	Stock of PUBLIC papers in a therapeutic class*	48.62	38.84	2.37	250.99

* All stocks are depreciated at a rate of 15% per year.

Table 2A
Distribution of PATENTS by selected Therapeutic Class and Selected Firms*

* Note that these firms are selected as illustrative of the pharmaceutical industry overall, and have not been chosen at randomly from our sample.

Table 2B
Distribution of PATENTS by Therapeutic Class & Location,
Merck Laboratories

Laboratory	Anti-psychotics	Antibiotics	Arrhythmia
Rahway, NJ	14%	27%	22%
West Point, PA	19%	4%	28%
Harlow, UK	64%	0%	14%
Dorval, CAN	14%	1%	48%

Table 3: Results from Poisson regressions, Y = PATENTS^

Variable	**Model 1** Coef. (StdErr) dY/dX	**Model 2** Coef. (StdErr) dY/dX	**Model 3** Coef. (StdErr) dY/dX	**Model 4** Coef. (StdErr) dY/dX	**Model 5** Coef. (StdErr) dY/dX	**Model 6** Coef. (StdErr) dY/dX	**Model 7** Coef. (StdErr) dY/dX	**Model 8** Coef. (StdErr) dY/dX
LOGDISC_IJ	0.106** (0.021) 0.201	0.111** (0.021) 0.205	0.112** (0.021) 0.207	0.116** (0.021) 0.212	0.087** (0.020) 0.150	0.109** (0.021) 0.199	0.111** (0.021) 0.202	0.000 (0.023) -0.001
LOGDISC_IJ_S	0.166** (0.020) 0.314	0.149** (0.020) 0.276	0.148** (0.020) 0.273	0.158** (0.020) 0.289	0.154** (0.020) 0.265	0.112** (0.020) 0.206	0.126** (0.020) 0.228	0.075** (0.022) 0.120
EXPOSURE_FLOW		0.021** (0.003) 0.039						
EXPOSURE_FLOW_PUB			0.088** (0.009) 0.162	0.117** (0.010) 0.213	0.111** (0.016) 0.191	0.154** (0.013) 0.282	0.179** (0.013) 0.324	0.168** (0.017) 0.268
EXPOSURE_FLOW_PRIV			-1.020** (0.209) -1.879	-0.988** (0.204) -1.806	-0.764** (0.216) -1.315	-1.016** (0.201) -1.858	-0.976** (0.196) -1.770	-0.688** (0.199) -1.099
N_CENTROIDS				-0.239** (0.044) -0.438			-0.259** (0.045) -0.470	-0.066 (0.047) -0.105
GLOBAL_PUB_STOCK					0.062** (0.004) 0.107			
GLOBAL_PRIV_STOCK					-0.597** (0.042) -1.028			
LOCAL_START_PUB						0.252** (0.039) 0.461	0.299** (0.040) 0.542	-0.133** (0.051) -0.212
LOCAL_START_PRIV						-2.031** (0.251) -3.715	-2.277** (0.259) -4.128	0.355 (0.328) 0.566
Intercept	-2.137** (0.336)	-2.669** (0.344)	-2.596** (0.341)	-2.623** (0.342)	-2.522** (0.343)	-2.792** (0.343)	-2.811** (0.343)	-2.024** (0.350)
N	704	704	704	704	704	691	691	691
Log L	-1895.7014	-1866.7618	-1850.7718	-1835.4602	-1734.7078	-1783.6405	-1766.6549	-1546.8752
Pseudo R-sq	0.2248	0.2367	0.2432	0.2495	0.2907	0.2609	0.2679	0.359
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Therapeutic class FEs								Yes+

^ dY/dX is computed at the mean of the independent variables. *=significant at 5%, **= significant at 1%. Dummy variables equal to 1 if LOGDISC_IJ and LOGDISC_IJ_S are undefined are also included in the regressions.

+ Class fixed effects are jointly significant. All but one are also individually significant at the 5% level.

Table 4: Results from Negative Binomial regressions, Y = PATENTS^

Variable	Model 1 Coef. (StdErr) dY/dX	Model 2 Coef. (StdErr) dY/dX	Model 3 Coef. (StdErr) dY/dX	Model 4 Coef. (StdErr) dY/dX	Model 5 Coef. (StdErr) dY/dX	Model 6 Coef. (StdErr) dY/dX	Model 7 Coef. (StdErr) dY/dX	Model 8 Coef. (StdErr) dY/dX
LOGDISC_IJ	0.111** (0.034) 0.209	0.107** (0.034) 0.202	0.111** (0.034) 0.208	0.115** (0.034) 0.215	0.077* (0.034) 0.140	0.095** (0.035) 0.179	0.101** (0.035) 0.184	0.002 (0.037) 0.004
LOGDISC_IJ_S	0.121** (0.029) 0.230	0.108** (0.029) 0.204	0.107** (0.029) 0.199	0.107** (0.029) 0.200	0.109** (0.028) 0.197	0.097** (0.029) 0.183	0.098** (0.029) 0.178	0.032 (0.032) 0.051
EXPOSURE_FLOW		0.016** (0.006) 0.030						
EXPOSURE_FLOW_PUB			0.071** (0.019) 0.133	0.092** (0.022) 0.172	0.084** (0.030) 0.152	0.106** (0.026) 0.201	0.121** (0.027) 0.219	0.114** (0.030) 0.180
EXPOSURE_FLOW_PRIV			-0.997** (0.363) -1.864	-1.007** (0.363) -1.879	-0.876* (0.361) -1.587	-1.036** (0.353) -1.955	-1.032** (0.353) -1.869	-0.796* (0.332) -1.251
N_CENTROIDS				-0.182 (0.094) -0.340			-0.187* (0.095) -0.338	-0.036 (0.089) -0.057
GLOBAL_PUB_STOCK					0.038** (0.008) 0.069			
GLOBAL_PRIV_STOCK					-0.356** (0.070) -0.645			
LOCAL_START_PUB						0.156* (0.070) 0.295	0.178* (0.070) 0.322	-0.106 (0.085) -0.166
LOCAL_START_PRIV						-1.250** (0.442) -2.359	-1.338** (0.441) -2.422	0.369 (0.538) 0.579
Intercept	-2.010** (0.368)	-2.388** (0.395)	-2.345** (0.388)	-2.350** (0.387)	-2.294** (0.387)	-2.420** (0.390)	-2.413** (0.388)	-1.699** (0.404)
N	704	704	704	704	704	691	691	691
Log L	-1413.44	-1409.28	-1405.02	-1403.17	-1391.93	-1383.95	-1382.05	-1315.63
Pseudo R-sq	0.0877	0.0904	0.0931	0.0943	0.1016	0.0955	0.0967	0.1402
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Therapeutic class FEs								Yes+

^ dY/dX is computed at the mean of the independent variables. *=significant at 5%, **= significant at 1%. Dummy variables equal to 1 if LOGDISC_IJ and LOGDISC_IJ_S are undefined are also included in the regressions.

+ Class fixed effects are jointly significant. All but one are also individually significant at the 5% level.

Figure 1
World Distribution of Cardiovascular Research, 1981-1990
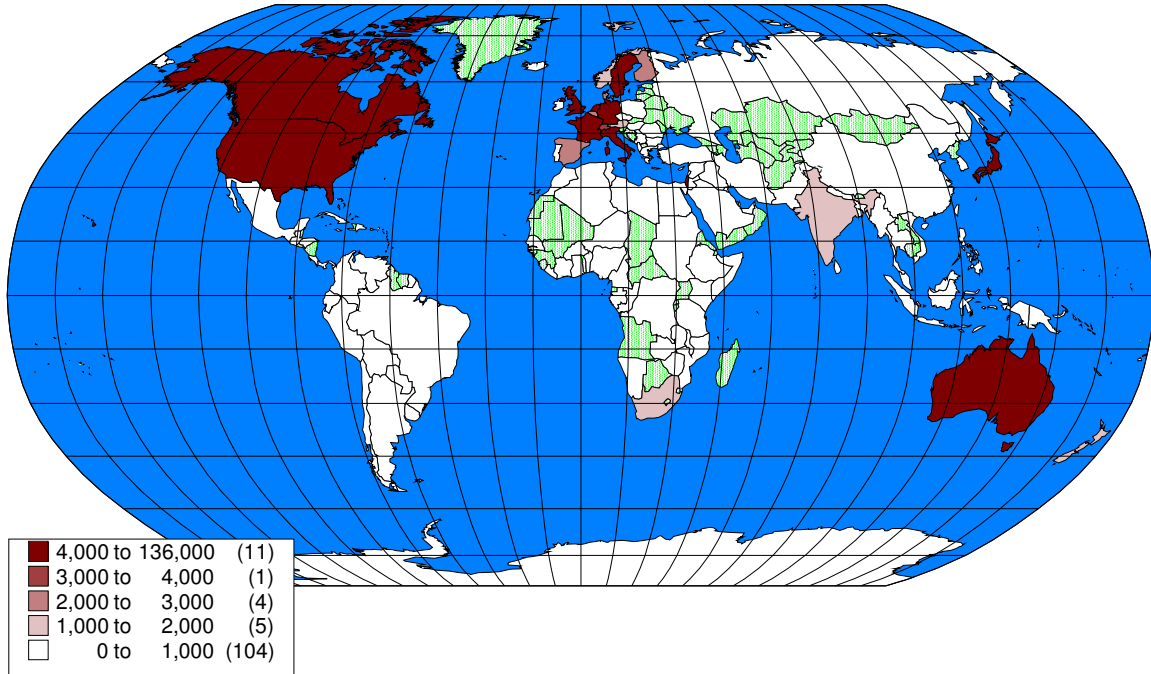
Distribution of Cardiovascular Research

4,000 to 136,000	(11)
3,000 to 4,000	(1)
2,000 to 3,000	(4)
1,000 to 2,000	(5)
0 to 1,000	(104)

Figure 2
World Distribution of Parisitology Research, 1981-1990

Distribution of Parisitology Research

50 to 9,890	(45)
40 to 50	(8)
30 to 40	(11)
20 to 30	(14)
0 to 20	(58)