NBER WORKING PAPER SERIES

UNIVERSITY RESEARCH, INDUSTRIAL R&D, AND THE ANCHOR TENANT HYPOTHESIS

Ajay Agrawal
Iain M. Cockburn

Working Paper 9212
http://www.nber.org/papers/w9212

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2002

We are grateful for helpful comments from Marc Busch, Wes Cohen, Alfonso Gambardella, John McHale, James MacKinnon, Mihkel Tombak, Andrew Toole, Arvids Ziedonis, and participants in the 2002 NBER Summer Institute and the 2002 Georgia Tech Roundtable for Engineering Research. We also appreciate the valuable efforts of Alex Oettl who provided excellent research assistance. This research was funded by the Queen's Center for Enterprise Development at Queen's University and the Boston University Systems Research Center. Their support is gratefully acknowledged. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

University Research, Industrial R&D, and the Anchor Tenant Hypothesis
Ajay Agrawal and Iain M. Cockburn
NBER Working Paper No. 9212
September 2002
JEL No. O3, R3, L3, L5

ABSTRACT

We examine geographic concentration, agglomeration, and co-location of university research and industrial R&D in three technological areas: medical imaging, neural networks, and signal processing. Using data on scientific publications and patents as indicators of university research and industrial R&D, we find strong evidence of geographic concentration in both activities at the level of MSAs. While evidence for agglomeration (in the sense of "excess" concentration relative to the size of MSAs and the size distribution of research labs) of research in these fields is mixed, we do find strong evidence of co-location of upstream and downstream activity. We view such co-located vertically connected activities as constituents of a "local innovation system," and these appear to vary markedly in their ability to convert local academic research into local commercial innovation. We develop and test the hypothesis that the presence of a large, local, R&D-intensive firm – an "anchor tenant" – enhances the productivity of local innovation systems by making local university research more likely to be absorbed by and to stimulate local industrial R&D. Presence of anchor tenant firms may be an important factor in stimulating both the demand and supply sides of local markets for innovation and may be an important channel for transmission of spillovers. While our empirical results are preliminary, they indicate that anchor tenant technology firms may be an economically important aspect of the institutional structure of local economies.

Ajay Agrawal
Queen's School of Business
Goodes Hall
Queen's University
Kingston, ON, Canada, K7L 3N6
aagrawal@business.queensu.ca

Iain M. Cockburn
School of Management, Boston University
595 Commonwealth Ave
Boston, MA 02215
and NBER
cockburn@bu.edu

1 Introduction

Much attention has been paid to the relationship between innovation, productivity, and geography. Such focus is understandable given the public policy implications and firm-location decisions that follow from regional variations in productivity related to the innovation process. At the same time, increased attention has been placed on the role of the university in the regional innovation system. This is at least in part due to the convergence of academic and industrial research interests in areas such as computer science, electrical engineering, and biotechnology.

In this paper, we set out to explore a set of questions concerning university research, industrial R&D, and their joint relationship with geography. To what extent are university research and, separately, industrial R&D, concentrated in specific technology areas? To what extent are these activities agglomerated? To what extent are they co-located? To what degree do regions vary in their productivity as measured by the industrial-R&D-to-university-research ratio? What might explain this variation?

We begin by examining the degree to which university research and, separately, industrial R&D associated with certain technical areas is concentrated and also agglomerated. We do this for three narrow technology areas in electrical engineering: medical imaging, neural networks, and signal processing.[1] While the concept of concentration is trivial, that of agglomeration is slightly more subtle. Agglomeration results in geographical clustering of activity and is usually thought to reflect externalities from localized knowledge spillovers and/or local "natural" advantage. Though agglomeration has been extensively documented for manufacturing activity (for example, Rosenthal and Strange, 2001), much less work has been done on industrial or academic innovative activity (Feldman, 1999). There is some

[1] Medical imaging technology facilitates the noninvasive generation of internal body images by employing techniques such as magnetic resonance imaging, ultrasound, nuclear medicine, and X-ray computed tomography. Neural networks are a form of multiprocessor computer systems based on collections of mathematical models that emulate some of the observed properties of biological nervous systems and draw on the analogies of adaptive biological learning. Signal processing technology enables the processing (e.g., filtering, compression, decompression) of signals (e.g., an analog electrical voltage or current, the digital output from the readout system of a compact disc player) and is used in a wide variety of applications such as mobile phones, multimedia computers, video recorders, hard disc drive controllers, and modems.

evidence that the geographic clustering of economic activity is stronger in research-intensive industries. Audretsch and Feldman (1996), and Jaffe *et al.* (1993) found that the "paper trail" of spillovers documented by patent citations is also significantly localized. Anecdotal evidence suggests that significant clusters may be present in areas such as pharmaceutical or biotech R&D[2] but there is little systematic evidence regarding its prevalence. While there remains little debate that many technologies are geographically concentrated, it does not follow that they are necessarily agglomerated. It is possible that agglomeration may be more apparent than real: a test for the presence of agglomeration effects must compare observed levels of concentration to those that would be expected to arise as the result of a random distribution of activity across geographic space, controlling for the size of the geographic unit and the size distribution of production units (Ellison and Glaeser, 1997).

We find little evidence of agglomeration of university engineering research in the three technologies under investigation. The apparent presence of clusters of academic activity in these technology fields in metropolitan statistical areas (MSAs) such as San Francisco-Oakland-San Jose, New York-Northern New Jersey-Long Island, Los Angeles-Riverside-Orange County, and Boston-Worcester-Lawrence is accounted for by higher levels of academic research in these MSAs and the size distribution of university labs. By contrast, for industrial R&D in these technologies, we do find some evidence of "excess" concentration. However, the quite low values which we obtain for the Ellison-Glaeser index are somewhat surprising: anecdotal evidence suggests a much larger agglomeration effect.

We then move on to explore the degree to which university research and commercial activity are co-located. Anecdotal evidence suggests that local innovation systems play a very important role in commercializing academic research, and our results suggest that these upstream-downstream activities are indeed co-located to a considerable extent. Our measures of academic research activity in these three technologies are strongly correlated at the regional level with measures of downstream industrial R&D activity, even after controlling for regional variations in size and economic activity. However, despite the positive correlation between

[2]See Prevezer (1997).

measures of these activities, we also note that there is significant variation across regions in the relationship between them. The productivity of local innovation systems varies quite substantially, and in the final section of the paper, we explore one hypothesis that may explain a significant portion of the variation. This is the “anchor tenant” hypothesis.

The classic “anchor tenant” is the large department store in a retail shopping mall that creates demand externalities for the other shops. Large department stores with a recognized name generate mall traffic that indirectly increases the sales of lesser-known stores. Pashigian and Gould (1998) examined the degree to which mall developers internalize these externalities by offering rent subsidies to anchor tenants and charging rent premiums to other mall tenants.

In our context, we define an anchor tenant as a large firm that is: 1) heavily engaged in R&D in general and 2) has at least minor absorptive capacity in a particular technology within a particular region. Anchor tenants may play a very important role in both creating and capturing externalities within local innovation systems. Anchor tenants create externalities by thickening markets and stimulating demand. They capture externalities by directly and indirectly increasing the absorptive capacity of the region for early-stage university-based research.

Our hypothesis is that the presence of an anchor tenant in an MSA enhances the regional innovation system such that local university research is more likely to be absorbed by and stimulate local industrial R&D. There are several reasons for thinking that this may be so. First, anchor tenants may be directly involved in the commercialization of university inventions. There are many examples of large, established firms working directly with universities in the context of collaborative research, co-supervising graduate students, sponsoring labs, licensing the rights to university inventions, recruiting graduate students, and hiring professors as consultants to directly leverage university research (Agrawal and Henderson, 2002; Agrawal, 2002).

Second, the anchor tenant may also indirectly stimulate innovative activity by enhancing both the supply and demand sides of the market for new technologies. Anchor tenants thicken factor markets such as labor, including both managerial and scientific. They also develop social networks with suppliers, buyers, and partners on which smaller firms can draw. The anchor tenant's demands for local resources such as intellectual-property legal counsel, technology-oriented marketing, and human resources services also thicken markets which benefit smaller firms.

Less often discussed is the effect of established firms on the demand side of technology markets. Anchor tenants are likely to enjoy economies of scope. Economies of scope that result from multiple product offerings enable anchor tenants to invest in the development of innovations with uncertain application that other firms may not be able to justify. To this end, anchor tenants play an important role by either directly absorbing local university research or absorbing further developed industrial R&D output from smaller firms.

Additionally, anchor tenants purchase products, licenses, consulting services, and perhaps entire companies. In other words, anchor tenants may facilitate a vibrant intermediate market between university research and large-scale production for end consumers. For example, many young technology firms survive on consulting revenues during their early years while they develop their first product (Roberts, 1991). The transaction costs associated with person-to-person interactions involved in consulting engagements strongly favor local clients. Thus, local anchor tenants that are consumers of local specialized consulting services may indirectly support industrial R&D in a particular technology area.

Anchor tenants also may utilize their economies of scope by bundling new innovations with existing products. So a young firm that develops a software innovation may license that innovation to the anchor tenant that bundles the software with its own hardware and distributes the product through its established distribution system. The high uncertainty associated with early-stage innovations, especially coupled with the high uncertainty associated with early-stage firms, favors local partners where frequent meetings with low transaction costs are possible. Thus, anchor tenants may play a critical, but indirect, role in stimulating industrial R&D that builds on local university research by stimulating downstream demand. In the final part of the empirical analysis, we specifically examine the *indirect* role of the anchor tenant by examining its effect on the industrial R&D output of the local economic "fringe."

Our findings support the anchor tenant hypothesis. We find a large positive and statistically significant effect of the presence of an anchor tenant on the relationship between academic engineering research in a region and corresponding downstream industrial R&D activity. Anchor tenant firms appear to play an important role in mediating research spillovers.

2 Literature

Our approach to exploring the questions articulated above builds on a rich literature that examines the economic and social factors at the nexus of innovation, productivity, and geography. In particular, our work draws from a subset of this literature that focuses on the role of the research university. This includes research on: 1) the characteristics of firms that utilize university spillovers, 2) the characteristics of universities that generate spillovers, 3) the characteristics of the channels through which university knowledge spills over, and 4) geography and university spillovers. We briefly review this prior work.

A variety of hypotheses have been proposed to explain the variance in ability of firms to utilize knowledge spillovers. Most of these focus on the capabilities and connectedness of the firm. Cohen and Levinthal (1989, 1990) introduce the concept of 'absorptive capacity' and argue that a firm's ability to apply university research for its own commercial gain is a function of its own investment in R&D. Cockburn and Henderson (1998) build on this notion, but add that the degree to which firms are "connected" to universities is also important for utilizing knowledge spillovers. Lim (2000) restructures the above two concepts and argues that the absorptive capacity of firms is primarily a function of its connectedness, of which its investment in in-house R&D is just one of several components.

Zucker *et al.* (1998) investigate the importance of connectedness to firms by examining their location decisions relative to star university scientists. Shane and Stuart (2002) study university start-up firms and examine the importance of connectedness, not with the scientific community, but rather with the venture capital community. Ziedonis (1999) does not consider connectedness, but instead examines the firm's related knowledge assets and its ability to evaluate external technology in terms of its likelihood of licensing a particular university technology as well as its likelihood of taking an option prior to licensing.

In addition to the connectedness and capabilities of firms that influence their ability to utilize university research, the characteristics of universities that influence spillovers from the supply side have been examined. Henderson *et al.* (1998) investigate the change in quality of university patents after the Bayh-Dole Act was passed in 1980. Thursby and Thursby (2002) develop a model to examine the extent to which the increase in university licensing is due to a fundamental change in the nature of research or just in the propensity to patent. Feldman *et al.* (2002) investigate the effect of licensing agreements involving equity rather than only cash payments. Jensen and Thursby (1998) examine the degree to which university inventions are so early-stage that they require the cooperation of the inventor to develop. Finally, Di Gregorio and Shane (2000) examine performance across university licensing offices and explore why some universities generate more new companies to exploit their intellectual property than others.

The channels through which knowledge is passed from universities to firms have been investigated from a number of perspectives. Cohen *et al.* (1998) and Cohen *et al.* (2002) both examine the relative importance of the complete set of transfer channels from the perspective of the knowledge recipient, namely firms. Agrawal and Henderson (2002) focus particularly on the comparison between patents and papers, but also analyze the relative importance of the complete set of transfer channels from the perspective of the knowledge creator, namely professors. Within the context of licensed patented inventions, Colyvas *et al.* (2002) examine the importance of transfer channels that complement patent licensing across different types of technologies. In addition, Shane (2002) investigates the question of when it is best for a university to license an invention back to the inventor by considering the effectiveness of different transfer channels, subject to the nature of the technology and its appropriability.

Finally, several studies focus on geography and knowledge spillovers. These studies measure knowledge inputs and associated outputs and examine their relationship across geographic space. The inputs and outputs considered vary from study to study, as does the geographic unit of analysis. Jaffe (1989) relates the input "federal research funding" to the output "new patents issued" at the state level. Jaffe *et al.* (1993) relate the input "original patents" to the output "patents that cite the original patents" at the city level. Audretsch and Feldman (1996) relate the input "local university research funding" to the output "local industry value-added" at the state level. Zucker *et al.* (1998) relate the input "number of local research stars" to the output "number of new local biotech firms" at the economic region level. Branstetter (2000) relates the input "scientific publications from the University of California" to the output "patents that cite those papers" at the the state level. Finally, Agrawal (2002) relates the input "hours of interaction with the MIT professor associated with a particular patented invention" to the output "likelihood or degree of success of commercializing that invention" and evaluates the impact of distance on this effect.

3 Variables and Data

For this study, we collected data on indicators of academic research activity and industrial R&D as well as a variety of control variables for the 268 US metropolitan statistical areas (MSAs) and consolidated metropolitan statistical areas (CMSAs) and the 25 Canadian census metropolitan areas (CMAs) - hereinafter collectively referred to as the "MSAs."[3] We removed 39 MSAs since their level of both academic and industrial electrical engineering research is *de minimis*, resulting in a working sample of 254 local economies.[4]

[3]While MSAs and CMAs are similar in spirit, they are defined slightly differently. The Canadian criterion requires that the urban core have a population of at least 100,000 for a metropolitan area to exist. In contrast, for the period 1990 to 2000, the United States had two criteria to determine whether or not a metropolitan area existed. In the United States, a metropolitan area exists where there is either a city of 50,000 or more inhabitants or a Census Bureau defined urban area, i.e., a population of at least 50,000 and a total metropolitan population of at least 100,000 (75,000 in New England). Thus, the Canadian approach is the more restrictive of the two.

[4]Specifically, MSAs were dropped from the sample if they had both fewer than five publications in IEEE journals per million inhabitants and fewer than 20 patents in electrical engineering per million inhabitants.

3.1 Industrial R&D (Patents)

We use patent counts to measure industrial R&D. Patent counts are generated by first creating a list of all patents that contain at least one US classification from a set of classifications associated with a particular technological area (e.g., medical imaging). The sets of US classifications used were created in consultation with electrical engineering professors who work on the technologies under investigation. The set of classifications associated with each technology is described in Appendix A. The queries are constrained to patents that have application dates between 1991 and 1997, inclusive. Patents are then assigned to MSAs by the address of the first inventor. It is important to note that the inventor MSA is often different from the assignee MSA. For example, many patents assigned to IBM Corp. in New York are invented by scientists and engineers located at labs in other MSAs.

There are three areas of concern associated with this measure of industrial R&D. First, not all industrial R&D results in patents. In fact, it has been well documented that for strategic reasons, many innovations generated from industrial R&D are protected by trade secret, other forms of intellectual property, or are not protected at all due to short product life cycles or difficulties associated with patenting (Cohen, 1998). This is particularly true for software for which there has been a substantial increase in the propensity to patent during the period under investigation.

Second, there is noise generated by our use of the classification system. US patent classifications do not exactly match the technology areas we are investigating. Most patents are assigned multiple classifications as their application spans several areas. However, the invention may be more related to some areas for which it is classified than others. As a result, the data suffers from both false positive and false negative results. In other words, the data contain patents that are not directly related to the technology under investigation and the data are missing patents that should be included. Third, not all patents represent the same level of economic importance.

To evaluate the seriousness of these concerns, we consider the degree of systematic bias by MSA that we might expect on any of these dimensions. For example, might some MSAs have a lower R&D-to-patent ratio than others? While we do not expect this to be a significant problem in general, results should be interpreted with caution and particular attention paid to the effect of MSAs that have a notably high propensity towards software development, such as Seattle and San Francisco.

We also do not expect a systematic bias in terms of classification-related noise by MSA. This is particularly true given the results we discuss later in the paper regarding the lack of agglomeration of commercialization activity in specific technological areas. Similarly, we do not expect a systematic bias in the economic importance of patents by MSA. However, this could be at least partly controlled for by weighting the importance of patents by citations, which is often done when using these measures. We plan to do this in the next draft of this paper.

The data used to generate this metric are collected from the United States Patent and Trademark Office (USPTO).

3.2 University Research (University Publications)

We use publication counts to measure university research activity. Specifically, we use author-MSA counts based on articles by university-based authors in particular journals. For example, to measure university research activity in the area of medical imaging, we begin by generating a list of all the articles that appeared in the journal *IEEE Transactions on Medical Imaging* during a particular time period (1991-1997). Then, we increment the MSA counter for the state associated with each university-based author. So, a paper by three authors, two of whom are from Boston-Worcester-Lawrence and one who is from Washington-Baltimore, will increment the Boston counter by two and the Washington counter by one. Also, each author is categorized by type: public, private, or university. Only university authors increment the MSA counters when measuring university research.

There are five areas of concern associated with this measure of university research. First, not all research is published. Agrawal and Henderson (2002) report that MIT electrical and mechanical engineering professors estimate that just less than 20% of the new knowledge generated by research at their labs utilized by industry is passed through the publishing channel. Other important channels of knowledge transfer include consulting, recruiting graduate students, research collaborations, and to a lesser extent, conferences, informal conversations, and patenting. Compared to Agrawal and Henderson, Cohen *et al.* (1998) report that almost the same fraction (20%) of U.S. manufacturing firms consider university publishing to be an important knowledge transfer channel for their industry. Thus, publishing only represents a fraction of overall university research output.

Second, not all published research on a particular technology is contained in the journals included in our analysis. For example, only a fraction of articles associated with medical imaging are published in the journal *IEEE Transactions on Medical Imaging*. Many other journals publish articles relevant to medical imaging. This is quickly verified by examining the citations in articles published in *IEEE Transactions on Medical Imaging*, which include references to dozens of other journals. However, other journals are not also included in the count because they also contain articles on topics outside of medical imaging.

Third, not all articles represent the same quality of research. Some articles are obviously more important than others and result from higher-quality research or longer-term projects.

A fourth issue is the degree to which articles are relevant for industrial application. Since the measure of university activity is being compared to a measure of industrial activity, an implicit assumption is that all research is equally relevant for commercial application. This assumption is clearly not true.

Finally, the author-MSA metric may cause concern. For example, consider the case of two articles of identical quality where one has a single author in Boston and the other has three coauthors in Boston. The first paper will increment the Boston paper counter by one, whereas the second paper will increment the Boston paper counter by three. So, coauthored papers have a greater influence on the research metric than single-authored papers.

Our first response to each of these concerns is to consider whether we might expect systematic bias on any of these dimensions across MSAs. For example, might university researchers in some regions be more likely than those in other regions to publish their results, rather than disseminate their results through other channels? Given the importance of publications as a nearly universal metric for research performance, this would seem unlikely. However, it might be the case that some regions have, on average, higher-quality engineering departments that attract faculty members who are more publishing-oriented than their colleagues elsewhere.

Similarly, are professors in some regions more likely than their colleagues in other regions to publish in the particular journals we include in our analysis, given that they publish their findings somewhere? We consulted two electrical engineering professors regarding which journals we should include in our analysis. We selected journals based on two criteria: 1) the journal is considered amongst the top journals that publish research on the topic under study and 2) all the articles included in the journal are related to the topic under study. We feel reasonably confident that the journals selected for this study fulfill these criteria. However, it may be the case that certain factors, such as the location of the journal editors during the period under study, introduce bias in terms of the frequency of articles that are published from particular regions.

There also is no reason to believe that the quality of articles by university researchers from particular regions is systematically higher than those from other regions. The editorial function of the journal should control for this, at least to some degree. However, one could argue that the top-ranked engineering departments are clustered in a small number of states, that research quality is highly correlated with department ranking, that journals accept papers that fall within a certain quality range, and thus that the quality of research published in a particular journal does indeed vary (within the range allowable by the particular publication) systematically by region. The electrical engineering professors we consulted with indicated that they do not believe this to be the case with the journals we examine in this study. However, we plan to control for this in the next draft of this paper by weighting papers by citation, which is a technique commonly used to control for "quality" by scholars who employ bibliometric data.

Perhaps our greatest area of concern with this measure comes from the potential bias associated with the relationship between the region and the degree to which research is applied rather than basic. University departments do sometimes tip towards a majority of either basic or applied research. However, we suspect that this problem is constrained by the editorial guidelines of the journals from which this data is generated, since we specifically selected journals that are reasonably applied.

Finally, the author-MSA metric offers the desirable characteristic of multiple counts per MSA per paper. Since we consider each publication to represent a research effort generating new knowledge, we count each region that has direct, local access to that effort and knowledge. Measures of local access to authors is important given our assumption about the tacit nature of knowledge associated with early-stage university research. So, in the case of papers with authors from different regions this metric attributes research activity to each region involved. Similarly, for papers that have multiple authors from the same region, the counter is incremented to reflect the number of authors. An alternative measure might be to only count first authors. This way, each paper would contribute equal weight. However, this would not reflect the degree to which coauthors facilitate increased access to and dissemination of tacit knowledge to industry.

The data used to generate this metric is collected from the Institute for Scientific Information's *Science Citation Index*. Data is collected from the journals *IEEE Transactions on Medical Imaging, IEEE Transactions on Neural Networks, and IEEE Transactions on Signal Processing* for analyses of medical imaging, neural networks, and signal processing, respectively. All articles from these journals that were published between the years 1991 through 1997, inclusive, are included in the analysis.

3.3 Anchor Tenant (Large R&D Capability, Absorptive Capacity in Particular Technology)

As described in the Introduction, we define an anchor tenant for a particular MSA and technology (e.g., Boston, medical imaging) as a firm that meets two conditions. First, the firm must have some absorptive capacity in the particular technology area. This is measured by the presence of at least one patent granted to the firm from the set of specified US classifications during the period under investigation (1991-1997 inclusive). Second, the firm must demonstrate that it is heavily involved in R&D in general. This is measured by the presence of at least one thousand patents granted to the firm, with any US classification, during the period under investigation. (We find similar results where this threshold is reduced to 500 patents.) A firm that meets both of these conditions is considered an anchor tenant.

For example, Texas Instruments, Inc. is an anchor tenant in the Dallas-Fort Worth metropolitan area in the technological areas of signal processing and neural networks. It satisfies the two conditions described above because it has conducted a great deal of R&D in general and at least some on signal processing and neural networks specifically. However, the same firm is not an anchor tenant in medical imaging as it has not conducted any R&D (by our measure) in this area and thus has no absorptive capacity related to this technology. Also, since our analysis is at the research facility level, it is possible for firms to be anchor tenants in multiple locations. For example, Motorola, Inc. is an anchor tenant in signal processing in both the Chicago-Gary-Kenosha and the Phoenix-Mesa metropolitan areas.

The idea of the "anchor tenant" effect is illustrated in a stylized fashion by four "matched pairs" of MSAs in Table I. These pairs of MSAs are similar in terms of size (population) and university research in the area of medical imaging (publications). However, their commercial R&D as measured by patents varies considerably. For example, the Los Angeles MSA is similar to the New York MSA in size and level of university research in medical imaging, but the New York MSA has two anchor tenants (IBM and Lucent) while Los Angeles has none in this technology area; New York has approximately double the number of industry patents as L.A. Similarly, the San Francisco and Boston MSAs are similar in size and level of university research in medical imaging, but San Francisco has two anchor tenants (Sun and HP) while Boston has none; again, the MSA with an anchor tenant has substantially more industry patents. The same pattern holds when comparing Minneapolis (3M) with Atlanta and Pittsburgh with Rochester (Eastman Kodak).

It is important to note that while the anchor tenant may play an important role in stimulating industrial research, that firm is not necessarily responsible for the majority of industrial research on that topic in the MSA under investigation. For example, Table II illustrates that the two anchor tenants in the New York MSA are only responsible for two of that region's 43 industrial patents in medical imaging. Similarly, Table III illustrates that the anchor tenant in the Minneapolis MSA is responsible for only three of that area's 15 industrial patents. When we test the anchor tenant hypothesis, we remove patents by the anchor tenant. The empirical analysis is described in detail below.

3.4 Control Variables

A variety of MSA level control variables are used throughout the analysis. These include population, personal income per capita, professional-scientific-technical services, general electrical engineering patents, general electrical engineering papers, all patents, and all papers from journals included in the ISI's Science Citation Index. Population data are from the U.S. Bureau of the Census[5] and Statistics Canada.[6] Personal income per capita and professional-scientific-technical services are also both from the Census Bureau [7] and Statistics Canada.[8] The Professional, Scientific, and Technical Services sector (sector 54) of the 1997 Economic Census covers establishments with payroll that specialize in performing professional, scientific, and technical activities for others. The Census Bureau states that "these activities require a high degree of expertise and training."

General electrical engineering patents are counted by including all US patents that have been designated international patent classifications within the set: G01-G12 or H01-H05. The general categories for these classifications are physics ("G") and electricity ("H"). General electrical engineering papers are counted by including all IEEE publications. The Institute for Electrical and Electronic Engineers (IEEE) is a non-profit, technical professional association.

[5] State and Metropolitan Area Data Book, 1997-98, Table B-1.

[6] 1996 Census of Canada, Profile Data, Ottawa, Canada.

[7] Gaquin, Derdre A. and DeBrandt, Katherine A., eds. *2000 County and City Extra: Annual Metro, City, and County Data Book, Ninth Edition*. Lanham, MD: Bernan Press, 2000. Table C, page 807-887. Data from the 1997 United States Economic Census, U.S. Census Bureau.

[8] Labour Force Historical Review (Statistics Canada data table), 1999. "Employment by Census Metropolitan Area, 3 month moving average." (averaged over 12 months to generate 1997 estimate). Accessed with Beyond 20/20 Professional Browser.

In 2002, the Institute had over 360,000 individual members in 150 countries, held over 300 major conferences per year, and claimed to produce 30 % of the world's published literature in electrical engineering, computers, and control technology.[9]

3.5 Descriptive Statistics

Table IV reports the descriptive statistics for variables constructed from these sources for each of the MSAs in our working sample.

While the range of the paper and patent counts in each technology area is quite large, their distribution is highly left-skewed. In medical imaging, 185 out 254 MSAs had no publications and 198 had no patents. Similarly, in signal processing, there were 154 MSAs with no publications and 108 with no patents, while in neural networks, 158 MSAs were without papers and 187 without patents. Signal processing is significantly "larger" than the other two technologies, averaging almost three times more papers and eight times more patents per MSA.

As can be seen in the Table, MSAs vary considerably in size, with population ranging from just over 82,000 (Pine Bluff, AR) to almost 20 million (New York–Northern New Jersey–Long Island.) Just under 9% of the sample are located in Canada. One measure of the technology-intensity of these local economies is the fraction of the population falling into the Professional, Scientific, and Technical (PST) category. PST workers averaged just under 2% of the population, with a maximum of 5% and a minimum of 0.3%. Another indicator is the volume of patenting. Patents issued to private-sector assignees in all technology classes averaged about 1000 per MSA, ranging from only 1 to more than 28,000. In per capita terms, patents in classes ranged eight per thousand inhabitants to less than 0.007. Just over 1/3 of these patents were in electrical engineering.

[9]http://www.ieee.org/about/

4 Empirical Testing

4.1 Agglomeration, Geographical Concentration, and the Location of Research Activity

Research activity in our three technology areas is not evenly dispersed across North America. In fact, counts of patents and papers in these areas are highly concentrated within a handful of MSAs. This concentration of activity can be seen in summary measures such as the "CR4." In each technology area, the top four MSAs account for about 42% of patents, and the four largest MSAs account for 24% of medical imaging papers, 27% of neural networks papers, and 27% of signal processing papers. A more broad-reaching measure of geographical concentration is the "locational Gini" coefficient, which measures the extent to which the distribution of activity across geographical units departs from a uniform allocation. Two versions of this measure are presented: the "raw" Gini coefficient, based on the share of each activity in total activity, and the "relative" Gini coefficient.[10] Table V reports the results of computing locational Gini coefficients for the paper and patent counts across MSAs.

The high values obtained for the raw Gini coefficient on paper counts by MSA indicate that academic research in all three technology areas is highly geographically concentrated. (If the activity is evenly distributed across geographical units, $G_L = 0$, while if all the activity is concentrated in a single geographical unit, $G_L \to 0.5$.) Interestingly, similar results are obtained when the Gini is recomputed relative to the distribution of total EE activity.[11] There are no substantive differences in the degree of geographic concentration across technologies or between papers and patents. The only markedly different results are obtained when we compare the distribution of electrical engineering in general. Compared to results for the narrower technology areas, academic publication activity in the broader field of electrical

[10]The locational Gini coefficient is reviewed in Krugman (1991) and Amiti (1998). The formula used here is $G_L \equiv \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{4n(n-1)\bar{x_i}}$, where for counts of activity X_i in $i = 1 \ldots n$ regions, x_i in the raw Gini is $X_i/\sum_i X_i$, and $\frac{X_i/\sum_i X_i}{Z_i/\sum_i Z_i}$ when the Gini is calculated relative to a broader activity Z_i.

[11]The relatively large numbers of MSAs with no measured activity may be biasing these calculations, but deleting observations with zero counts reduced these coefficients by only a few percent.

engineering is somewhat less geographically concentrated relative to the distribution of all scientific publishing, and commercial patenting activity in electrical engineering relative to total patenting is much less concentrated.

Of course, simply observing this geographic concentration is not an indication that the location of research is necessarily a reflection of agglomeration driven by localized spillovers or natural advantage. To infer agglomeration from the geographic distribution of activity, its degree of concentration must be compared to that which would arise from a random distribution of production units across geographic areas. One way to do this is by calculating the Ellison-Glaeser (1997) index, which controls for variation in the size distribution of production units and in the size of geographic areas as well as compares the actual distribution of activity with a benchmark corresponding to a "dartboard" allocation of production units across states.

Table VI presents values of the Ellison-Glaeser index computed for each measure of research activity. The index is given by the formula

$$\gamma \equiv \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - (1 - \sum_{i=1}^{M} {x_i}^2)H}{(1 - \sum_{i=1}^{M} {x_i}^2)(1 - H)} \tag{1}$$

where, for papers, s_i is the share of each MSA in publications in the technology area, x_i is each MSA's share of total academic publications, and H is the Herfindahl index of publications in the technology area across universities. For patents, s_i is the share of each MSA in patents in that technology area, x_i is the share of each state in total patents, and H is the Herfindahl index of patents in the technology across assignee-MSA units.

Interestingly, we find only mixed evidence of agglomeration in any of the three technology areas. In the first panel of Table VI, γ is computed relative to electrical engineering, that is to say, the x_i's are MSA shares in total academic electrical engineering publications, or total electrical engineering patents. While there appears to be more agglomeration in commercial R&D activity than in academic research, the γs are all quite small. Somewhat different results are obtained when the γs are computed relative to total scientific or R&D activity in all areas. As can be seen in the second panel of Table VI, the γs for patents are significantly larger, particularly for neural networks and signal processing, where they are almost 30 times the value in the first panel. Though it is difficult to interpret the index values other than relative to a salient benchmark, Ellison and Glaeser suggested a cutoff of $\gamma < 0.02$ to describe US manufacturing industries as "not very concentrated," and we infer from these results that some degree of "excess" concentration of R&D activity within MSAs is present.

4.2 Are Papers and Patents Co-located?

Although we find mixed evidence for agglomeration of publication activity or patenting activity in these technologies, this does not mean that publication and patenting are not co-located.[12] Indeed, if our notion of localized innovation systems is correct, then a strong statistical relationship should exist between the level of publications in a given technological area produced in an MSA and the level of patents in the technological area which originate in that MSA.

One piece of evidence for co-location is the raw correlation between papers and patents across MSAs, which is quite high: 0.67 for medical imaging, 0.52 for neural networks, and 0.77 for signal processing (Table VII.) These correlations are robust to the exclusion of outliers, and non-parametric measures of association derived from the cross-tabulation of counts show a positive and statistically significant relationship. But this measure confounds the effect of the size of the MSA on both variables, so a better descriptive parameter is the coefficient on papers in a regression of patents on to papers and a control for the size of the MSA.

Tables VIII gives results from regressing patents on to papers. The univariate regression results simply restate the correlation coefficients. But even when we control for size using PST LABOR, in each of the three technologies we get a positive and strongly significant coefficient on papers. The same result holds using various other size controls such as population or the total number of IEEE publications or electrical engineering patents originating

[12] Co-location is different from co-agglomeration. Along with their index of agglomeration, Ellison-Glaeser propose an index of co-agglomeration which accounts for size differences across geographic and production units. As with the univariate indexes, we find mixed evidence for significant "excess" co-location of papers and patents in each technology area: relative to electrical engineering, we obtained values of Ellison-Glaeser's γ^c of 0.008 for medical imaging, 0.004 for neural networks, and 0.007 for signal processing; relative to all science, we obtained values of 0.002 medical imaging, -0.009 for neural networks, and 0.053 for signal processing. Again, these numbers are difficult to interpret without an appropriate benchmark.

from the MSA. Taken at face value, these results imply that there is a very strong co-location effect: local patents in a technological area are strongly associated with the presence of upstream academic science. Here, and in the additional regressions reported below, we place no structural interpretation on the results. Though there are good reasons to believe that papers "cause" patents in the sense that downstream industrial R&D activity relies on upstream science, it is quite possible that causation runs in the opposite direction. We have not specified a production function technology for R&D nor any assumptions about the behavior of actors in this process. Rather, the regressions are presented as descriptive analyses of reduced form associations in the data.

These OLS regressions are estimated using a simple linear functional form; given the disparity in size over the sample of MSAs, there is likely to be a significant amount of heteroscedasticity. White's test statistic confirms this hypothesis, rejecting the null of i.i.d. errors at $P > 0.0001$. The presence of many zeroes in both the dependent variable and the main explanatory variable prevents us from using a log-log specification, which would be less vulnerable to this problem. We therefore re-estimated some equations using weighted least squares to account for heteroscedasticity under the assumption that the variance of the error term is inversely proportional to population. This has a relatively small impact on the standard errors relative to the unweighted regressions.[13]

One important source of specification error arises from the measurement properties of the dependent variables in these regressions. As counts of patents, these variables take on only non-negative integer values, suggesting that a Poisson-type model is appropriate. Table IX presents results from re-estimating these models using Poisson regressions. The functional form is

$$E[PATENTS] = exp(\alpha + \beta PAPERS) \tag{2}$$

Note that since PAPERS is zero for many observations, it is not possible to estimate the standard log-log functional form, and β cannot therefore be read directly as an elasticity. As in the OLS regressions, we again find a positive and strongly significant coefficient on PAPERS

[13]However, computing robust standard errors using Stata's Huber-White estimator resulted in substantial increases, suggesting that the linear model is seriously misspecified.

in all three technologies. However, this is driven largely by confounding with the size of MSAs; when ln(PST LABOR) is included in the regression, the PAPERS variable is knocked out. Moreover, these data are badly over-dispersed relative to the Poisson distribution: the χ^2 goodness-of-fit statistic is very large in all cases, rejecting the null of $mean = variance$ at $P > 0.0001$.

One option for addressing over-dispersion is to use negative binomial regression, effectively adding a random effect to the model. As can be seen in Table X, while the coefficient estimates are broadly similar, the standard errors increase by an order of magnitude. Using this specification, PAPERS is only significant for the case of signal processing. However, a closer look at the data suggests that the negative binomial model may not be appropriate. The large number of zero patent counts suggests an alternative reason why the $mean = variance$ property of the Poisson distribution is violated here; many MSAs appear to have no activity in these technology areas, and therefore cannot be expected to generate patents. If this is the case, then a different model is appropriate. Here we use the ZIP ("Zero-Inflated Poisson") model proposed by Lambert (1992).[14] In this model, there are two unobserved states of the world: Regime 1) counts greater than zero are very rarely observed and Regime 2) counts are generated by a standard Poisson process in which zeroes can occur, but counts of one or more are also likely to be observed. The probability of being in one regime or the other can be modeled as logit or probit, with counts in the latter regime following the usual Poisson density function. The likelihood function is built up from two components, the Poisson equation and the "inflation" equation, and two sets of coefficients are estimated, one for the covariates in the Poisson part of the model and one for the logit/probit "inflation" equation.

Results from estimating the ZIP model are presented in Table XI. In each case, the "inflation" equation is assumed to be logit, with ln(population), share of EE in total patents, and a dummy for NO IEEE PAPERS as explanatory variables. In two out of three technologies, we again find a positive and significant coefficient on papers.

[14] See also Greene (1994).

4.3 Anchor Tenants and Productivity

Even after controlling for size, there is significant dispersion about the regression line of patents on papers. (See Figures 1 to 3.) One way to interpret this dispersion is in terms of ability of MSAs' local innovation systems to convert local science into local industrial innovation. Substantial variation in the patent/paper ratio even after controlling for size suggests that other characteristics of MSAs may be important determinants of the productivity of local innovation systems. Indeed, our core hypothesis in this study is that the presence of anchor tenants is a significant factor driving productivity. A simple test of this hypothesis is to include a measure of anchor tenants in the regressions of patents on to papers. Results for the OLS model are presented in Table XII. The measure of the presence of anchor tenants is ANCHORS, an indicator variable which =1 when there is at least one large anchor tenant in the MSA. ("Large" anchor tenants were issued at least 1000 patents during the sample period.) Essentially similar results were obtained using a measure based on the presence of medium-sized anchor tenants (500 patents) and using a direct count of anchor tenants instead of a dummy variable. As before, we control for the size of the MSA with PST LABOR and use weighted least squares to control for heteroscedasticity. Note that patents assigned to the anchor firm(s) are netted out from the dependent variable, so that the model should be thought of as referring to the relationship between university research and the patenting of the "fringe" of non-anchors. [15]

The results are striking: in all three technology areas, the coefficient on the interaction term is positive and strongly significant. For medical imaging and signal processing in MSAs which have an anchor tenant, the coefficient on PAPERS is roughly two to three times larger. For neural networks, the coefficient on PAPERS goes from negative and weakly significant to positive and significant when an anchor tenant is present.

The same flavor of results is obtained from estimating Poisson and Negative Binomial

[15] Anchor firms account for a varying fraction of total patenting in each technology and MSA. On average, this fraction is quite small (about 2%) but in a small number of MSAs it is much larger (over 90%). Netting out patents by anchor firms is important to avoid inducing a spurious correlation between the dependent variable and the anchor variable, but, as can be seen in the correlations reported in Table VII, the two counts are effectively very similar.

versions of this model, though, as argued above, there is reason to believe that these two models are misspecified. Given the multiplicative model and problems in handling zeroes in the PAPERS variable, ANCHOR enters the regression without an interaction term. Again, we find evidence for a positive impact of anchor tenants on the rate at which academic research is converted into patents in the local economy. The coefficient on the ANCHOR dummy is positive for all three technologies, though the standard errors are unstable. Our preferred specification is the ZIP model (see Table XIII), and here we find positive and strongly significant effects of the presence of anchor tenants. The impact is considerable: at the mean, the marginal impact of one more paper on the expected number of patents in medical imaging is 1.55 times larger in the presence of an anchor tenant, 2.7 times larger in neural networks, and 1.64 times larger in signal processing.

5 Conclusions

Important connections between university research and industrial R&D exist, but these are subtle and quite difficult to capture empirically. Anecdotal evidence suggests an important role for universities in generating clusters of innovative "spin-off' companies, although econometric evidence on the presence of localized knowledge spillovers from universities is mixed. The results reported by, for example, Jaffe (1989), Acs, Audretsch, and Feldman (1992), and Anselin, Varga, and Acs (1997) provide some evidence for positive externalities generated by university research operating over MSA-scale distances, but the estimated magnitude of these effects is quite sensitive to the level of aggregation over industries and technologies, as well as the definition of the local area. These difficulties in identifying a large, uniform, and real effect of academic research operating at this geographical scale suggest that much closer attention to measurement issues, as well as to the mechanisms and institutions through which these spillovers are transmitted, is needed. Our findings here support this view and prompt further close scrutiny of the phenomenon of clustering of R&D activities and of the institutional structure of local innovation systems.

Based on the distribution of publications and patents across North American MSAs in three sub-areas of electrical engineering, we find evidence of strong geographic concentration of research. But it does not necessarily follow from this that private-sector innovative activity is clustered in the sense of being over-concentrated relative to: 1) the geographic distribution of research activity and 2) the size distribution of labs. Controlling for these factors, we find some significant departures in the distribution of private-sector innovative activity from that predicted by a "dartboard" model of location. But we do not find any evidence for "excess" geographic concentration of university research in these three technologies, suggesting that the location of academic research may be driven by factors other than horizontal spillovers or Marshallian natural advantage.

We do find strong evidence for the co-location of downstream industrial R&D with upstream university research at the level of the MSA. While we hesitate to draw any conclusions about a causal relationship between academic research and industrial R&D, the degree of geographical association between these activities suggests a substantial localized component of vertical knowledge spillovers. Interestingly, the magnitude of this effect appears to be strongly mediated by the presence of anchor tenant firms in the local economy. Again, there are obvious and potentially serious endogeneity issues which we have not yet addressed, but the size and persistence of this effect in our regression results suggest an economically significant phenomenon.

Thus, the anchor tenant hypothesis may have implications for policy and strategy. Policymakers are often involved in the courtship process when large firms are investigating potential sites for a new location and accordingly may offer substantial incentives to attract such firms. In these negotiations, local governments need to know how much a large firm is worth to the local economy; thinking about the role of anchor tenants in mediating spillovers and shaping natural advantage may provide a useful framework for analysis. In this context it may be helpful to consider the number and significance of industries and technologies for which the firm in question might be an anchor tenant in the sense we have identified here, as well as to evaluate the extent to which the local universities produce research that is complementary to that firm's technologies.

Equally, externalities associated with anchor tenants have implications for corporate strategy. For example, when firms explore sites for large new R&D facilities, they may consider the extent to which they are likely to serve as an anchor tenant in the receiving local economy and the effectiveness of strategies for capturing externalities generated by their presence. Apart from taking advantage of local governments' willingness to offer incentives to locate in their jurisdiction, anchor tenants may be able to capture externalities through taking advance equity stakes in "fringe" firms, or through other private contractual arrangements. At the same time, smaller firms may realize substantial benefits from considering the presence of anchor tenants when making their location decisions.

Of course, while our results suggest that anchor tenant firms may be an important aspect of the institutional structure of local innovation systems, it is undoubtedly premature to speculate about specific policy implications. Much work remains to further investigate this phenomenon.

References

[1] Acs, Z., D. Audretsch, and M. Feldman (1992). "Real Effects of Academic Research: A Comment." *American Economic Review*, 82, pp. 262-367.

[2] Agrawal A. (2002). "Technology Acquisition Strategy for Early-Stage Innovations: Measuring the Importance of Engaging the Inventor." Mimeo, Queen's University.

[3] Agrawal, A. and R. Henderson (2002). "Putting Patents in Context: Exploring Knowledge Transfer from MIT." *Management Science*, Vol. 48, No. 1, pp. 44-60.

[4] Amiti, Mary (1998). "New Trade Theories and Industrial Location in the EU: A Survey of Evidence." *Oxford Review of Economic Policy*, 14(2), pp. 45-53.

[5] Anselin, L., A. Varga, and Z. Acs (1997). "Local Geographical Spillovers Between University Research and High Technology Innovations." *Journal of Urban Economics*, 42, pp. 153-182.

[6] Audrestch, D. and M. Feldman (1996). "R&D Spillovers and the Geography of Innovation and Production." *American Economic Review*, Vol. 86, No. 3, pp. 630-640.

[7] Blumenthal, David (1996). "Relationships between Academic Institutions and Industry in the Life Sciences - An Industry Survey." *New England Journal of Medicine*, Vol. 334, No. 6, pp. 368-73.

[8] Branstetter, L. (2000). "Measuring the Link Between Academic Science and Industrial Innovation: The Case of California's Research Universities." Paper presented at the NBER Summer Institute.

[9] Cockburn, I., R. Henderson, and S. Stern (2000). "Untangling the Origins of Competitive Advantage." *Strategic Management Journal*, Vol. 21, pp. 1123-1145.

[10] Cockburn, I., R. Henderson, and S. Stern (2001). "Balancing Incentives: The Tension between Basic and Applied Research." Mimeo, MIT.

[11] Cockburn, I. and R. Henderson (1998). "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery." *Journal of Industrial Economics*, Vol. XLVI, No. 2. pp. 157-182.

[12] Cohen W. and D. Levinthal (1989). “Innovation and Learning: The Two Faces of R&D.” *The Economic Journal*, Vol. 99, pp. 569-596.

[13] Cohen W. and D. Levinthal (1990). “Absorptive Capacity: A New Perspective on Learning and Innovation.” *Administrative Sciences Quarterly*, Vol. 35, pp. 128-152.

[14] Cohen, W., R. Florida, L. Randazzese, and J. Walsh (1998) “Industry and the Academy: Uneasy Partners in the Cause of Technological Advance.” *Challenges to the Research University,* ed. R. Noll. Brookings Institution, Washington, D.C.

[15] Cohen, W., R. Nelson, and J. Walsh (2002). “Links and Impacts: The Influence of Public Research on Industrial R&D.” *Management Science*, Vol. 48, No. 1, pp. 1-23.

[16] Colyvas, J., M. Crow, A. Geligns, A. Mazzoleni, R. Nelson, N. Rosenberg, and B. Sampat (2002). “How Do University Inventions Get Into Practice?” *Management Science*, Vol. 48, No. 1, pp. 61-72.

[17] Di Gregorio, D. and S. Shane (2000). “Why Do Some Universities Generate More Start-Ups than Others?” Mimeo, University of Maryland.

[18] Ellison, G. and E. Glaeser (1997). “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach.” *Journal of Political Economy*, Vol. 105, No. 5, pp. 889-927.

[19] Feldman, M. (1999). “The New Economics of Innovation, Spillover and Agglomeration: Review of Empirical Studies.” *Economics of Innovation and New Technology*, Vol. 8, No. 1-2, pp. 5-25.

[20] Feldman, M., I. Feller, J. Bercovitz, and R. Burton (2002). “Equity and the Technology Transfer Strategies of American Research Universities.” *Management Science*, Vol. 48, No. 1, pp. 105-121.

[21] Greene, W. (1994). “Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models.” NYU Department of Economics WP No. EC-94-10, New York, NY.

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

[23] Henderson R., A. Jaffe and M. Trajtenberg (1998). "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988." *Review of Economics and Statistics*, Vol. 80, No. 1, February, pp. 119-127.

[24] Jaffe, A. (1989). "Real Effects of Academic Research." *American Economic Review*, 79, pp. 957-970.

[25] Jaffe, A., M. Fogarty, and B. Banks (1997). "Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation." NBER Working Paper No. 6044, Cambridge, MA.

[26] Jaffe, A. and M. Trajtenberg (1996). "Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries." NBER Working Paper No. 5712, Cambridge, MA.

[27] Jaffe, A., Henderson, R., and M. Trajtenberg (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*, 108, pp. 577-598.

[28] Jensen, R. and M. Thursby (1998). "Proofs and Prototypes for Sale: The Tale of University Licensing." NBER Working Paper No. 6698, Cambridge, MA.

[29] Krugman, Paul (1991). *Geography and Trade.* MIT Press, Cambridge, MA.

[30] Lambert, Diane (1992). "Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing." *Technometrics*, Vol. 34, No. 1, pp. 1-14.

[31] Lim, K. (2000). "The Many Faces of Absorptive Capacity: Spillovers of Copper Interconnect Technology for Semiconductor Chips." Mimeo, MIT.

[32] Mansfield, E. (1995). "Academic Research Underlying Industrial Innovations: Sources, Characteristics and Financing." *The Review of Economics and Statistics*, 77, pp. 55-65.

[33] Mowery, D.C., R.R. Nelson, B. Sampat, and A.A. Ziedonis (2002). "Academic Patent Quality and Quantity Before and After the Bayh-Dole Act in the United States." Forthcoming in *Research Policy.*

[34] Pashigian, P. and E. Gould (1998). "Internalizing Externalities: The Pricing of Space in Shopping Malls." *Journal of Law and Economics*, Vol. XLI, pp. 115-142.

[35] Prevezer, M. (1997). "The Dynamics of Industrial Clustering in Biotechnology." *Small Business Economics*, Vol, 9, No. 3, pp. 255-271.

[36] Rosenthal, S. and Strange, W. (2001). "The Determinants of Agglomeration." *Journal of Urban Economics*, 50(2), pp. 191-229.

[37] Shane, S. (2002). "Selling University Technology: Patterns From MIT." *Management Science*, Vol. 48, No. 1, pp. 122-137.

[38] Shane, S., and T. Stuart (2002). "Organizational Endowments and the Performance of University Start-ups." *Management Science*, Vol. 48, No. 1, pp. 154-170.

[39] Thursby, J. and M. Thursby (2002). "Who is Selling the Ivory Tower? Sources of Growth in University Licensing." *Management Science*, Vol. 48, No. 1, pp. 90-104.

[40] Ziedonis, A. (1999). "Inward Technology Transfer by Firms: The Case of University Technology Licenses." Mimeo, University of California, Berkeley.

[41] Zucker, L., Darby, M., and J. Armstrong (1998). "Intellectual Capital and the Firm: The Technology of Geographically Localized Knowledge Spillovers." *Economic Inquiry*, Vol. 36, pp. 65-86.

[42] Zucker, L., M. Darby, and M. Brewer (1998). "Intellectual Capital and the Birth of U.S. Biotechnology Enterprises." *American Economic Review*, Vol. 88, pp. 290-306.

[43] Zucker, L., M. Darby, and J. Armstrong (2002). "Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology." *Management Science*, Vol. 48, No. 1, pp. 138-153.

Table I: Matched Pairs: Example from Medical Imaging

MSA	Population	University Papers	Industry Patents	Anchor Tenant
L.A. Riverside Orange County	15608886	55	21	None
New York Northern NJ Long Isl.	19876488	42	43	IBM, Lucent
San Francisco Oakland San Jose	6700753	30	86	Sun, HP
Boston Worcester Lawrence	5827654	28	42	None
Minneapolis St. Paul	2792137	14	15	3M
Atlanta	3627184	11	6	None
Pittsburgh	2361019	7	1	None
Rochester	1086082	6	34	Eastman Kodak

Table II: Matched Pairs: Los Angeles versus New York

	Total Patents	Company MI Patents	Anchor
Los Angeles (21 total MI patents)			
Capistrano Labs Inc.	1	1	
Cardiovascular Imaging Systems Inc.	4	4	
Cordis Webster Inc.	19	1	
Imagyn Medical Technologies Inc.	4	1	
Integrated Medical Systems Inc.	1	1	
International Remote Imaging Systems Inc.	11	5	
Johnson & Johnson Medical Inc.	27	1	
Logicon Inc.	1	1	
Northrop Grumman Corp.	184	1	
Sonus Pharmaceuticals Inc.	1	1	
TOA Medical Electronics Co. Ltd.	2	1	
Vivorx Pharmaceuticals Inc.	16	3	
New York (43 total MI patents)			
Biosense Inc.	6	1	
Center for Laboratory Technology Inc.	1	1	
Ciba Geigy Corp.	178	1	
Cytometrics Inc.	1	1	
Echocath Inc.	5	1	
IBM Corp.	3281	1	X
Lucent Technologies Inc.	1632	1	X
Mobil Oil Corp.	68	1	
Neoromedical Systems Inc.	6	5	
Ortho Diagnostic Systems Inc.	11	1	
Phillips Electronics NA Corp.	257	5	
Sarnoff Corp.	174	5	
Schick Technologies Inc.	1	1	
Siemens Corporate Research Inc.	104	17	
Trex Medical Corp.	5	1	

Table III: Matched Pairs: Minneapolis versus Atlanta

	Total Patents	Company MI Patents	Anchor
Minneapolis (15 total MI patents)			
Clarus Medical Systems Inc.	10	2	
Eastman Kodak Co.	11	1	
Imation Corp.	105	2	
Insight Medical Systems Inc.	1	1	
Medtronic Inc.	470	1	
Micro Medical Devices Inc.	2	1	
Minnesota Mining and Manufacturing Co.	2092	3	X
Picker International Inc.	2	2	
Shturman Cardiology Systems Inc.	15	2	
Atlanta (6 total MI patents)			
General Electric Co.	6	1	
Georgia Tech Research Corp.	129	2	
Minolta-QMS Inc.	1	1	
North American Phillips Corp.	3	2	

Table IV: Descriptive Statistics ($N = 254$)

Variable	Mean	Std Dev	Min	Max
Medical Imaging				
Papers	2.567	7.225	0	55
Patents	2.323	9.567	0	86
Patents (excl. anchors)	2.008	8.947	0	85
No. of anchors (> 1000)	0.031	0.196	0	2
No. of anchors (> 500)	0.039	0.232	0	2
No. of anchors (> 100)	0.122	0.552	0	5
Neural Network				
Papers	2.327	6.581	0	65
Patents	2.287	9.168	0	98
Patents (excl. anchors)	1.831	7.413	0	74
No. of anchors (> 1000)	0.067	0.307	0	2
No. of anchors (> 500)	0.142	0.599	0	6
No. of anchors (> 100)	0.315	1.211	0	13
Signal Processing				
Papers	7.205	19.171	0	154
Patents	18.618	72.941	0	754
Patents (excl. anchors)	15.106	59.354	0	641
No. of anchors (> 1000)	0.079	0.358	0	3
No. of anchors (> 500)	0.169	0.769	0	9
No. of anchors (> 100)	0.602	2.200	0	24
General				
Canadian	0.087	0.282	0	1
Papers (all IEEE)	133.784	324.467	0	2268
Patents (all EE)	387.228	1328.765	0	14227
Papers (all in ISI)	8392.484	19756.09	0	165664
Patents (all in USPTO)	1036.563	2945.978	1	28316
Population	877316.4	1943850	82024	1.99e+07
Income per capita (000's)	22.750	3.986	2.937	38.772
PST labor (000's)	21.332	59.219	0.503	569.807

Table V: Gini Coefficients

	Papers	Patents
Medical Imaging		
Raw locational coeff.	0.443	0.469
Locational coeff. relative to EE	0.440	0.456
Neural Networks		
Raw locational coeff.	0.424	0.458
Locational coeff. relative to EE	0.451	0.454
Signal Processing		
Raw locational coeff.	0.427	0.449
Locational coeff. relative to EE	0.384	0.363
EE locational coeff. relative to total	0.342	0.195

$N = 254$

Table VI: Ellison-Glaeser Indexes

1. Relative to Electrical Engineering Research				
	$\sum_i x_i^2$	$\sum_i (s_i - x_i)^2$	H	γ
Papers				
Medical Imaging	0.027	0.010	0.019	-0.009
Neural Networks	0.027	0.009	0.018	-0.009
Signal Processing	0.027	0.004	0.015	-0.012
Patents				
Medical Imaging	0.050	0.024	0.019	0.006
Neural Networks	0.050	0.010	0.008	0.002
Signal Processing	0.050	0.005	0.004	0.001
2. Relative to All Research Activity				
	$\sum_i x_i^2$	$\sum_i (s_i - x_i)^2$	H	γ
Papers				
Medical Imaging	0.026	0.000	0.019	-0.011
Neural Networks	0.026	0.013	0.018	-0.006
Signal Processing	0.026	0.005	0.015	-0.010
Patents				
Medical Imaging	0.036	0.035	0.019	0.018
Neural Networks	0.036	0.035	0.008	0.028
Signal Processing	0.036	0.034	0.004	0.031

$N = 254$

Table VII: Correlations Between Patents and Papers

Medical Imaging			
	Patents (all)	Patents (excl. anchor)	Papers
Patents (all)	1.000		
Patents (excl. anchor)	0.945	1.000	
Papers	0.581	0.606	1.000
Neural Networks			
	Patents (all)	Patents (excl. anchor)	Papers
Patents (all)	1.000		
Patents (excl. anchor)	0.977	1.000	
Papers	0.520	0.522	1.000
Signal Processing			
	Patents (all)	Patents (excl. anchor)	Papers
Patents (all)	1.000		
Patents (excl. anchor)	0.983	1.000	
Papers	0.761	0.766	1.000

$N = 254$

Table VIII: Regression Results: OLS

	Medical Imaging		Neural Networks		Signal Processing	
	Patents	Patents	Patents	Patents	Patents	Patents
Papers	0.769***	0.424***	0.724***	-0.277***	2.894***	1.244***
	(0.068)	(0.110)	(0.075)	(0.089)	(0.156)	(0.261)
PST Labor		0.053***		0.142***		0.635***
		(0.013)		(0.010)		(0.085)
Constant	0.349	0.113	0.603	-0.102	-2.232	-3.878
	(0.520)	(0.509)	(0.522)	(0.392)	(3.181)	(2.889)
$Adj R^2$	0.334	0.370	0.267	0.595	0.577	0.653

*significant at the 0.1 level, ** 0.05, ***0.01.

$N = 254$, standard errors in parentheses

Table IX: Regression Results: Poisson

	Medical Imaging		Neural Networks		Signal Processing	
	Patents	Patents	Patents	Patents	Patents	Patents
Papers	0.083***	0.002	0.063***	-0.006**	0.031***	0.004***
	(0.002)	(0.004)	(0.0002)	(0.003)	(0.000)	(0.000)
$ln(PSTLabor)$		1.025***		1.082***		0.987***
		(0.046)		(0.039)		(0.017)
Constant	0.159***	-2.369***	0.421***	-2.505***	2.156***	-0.322***
	(0.056)	(0.161)	(0.050)	(0.154)	(0.020)	(0.059)
$PseudoR^2$	0.359	0.558	0.248	0.594	0.547	0.762

*significant at the 0.1 level, ** 0.05, ***0.01.

$N = 254$, standard errors in parentheses

Table X: Regression Results: Negative Binomial

	Medical Imaging Patents	Neural Networks Patents	Signal Processing Patents
Papers	-0.007	0.041	0.019***
	(0.030)	(0.037)	(0.007)
$ln(PSTLabor)$	1.661***	0.826***	0.952***
	(0.217)	(0.130)	(0.079)
Constant	-4.378***	-1.964***	-0.587***
	(0.547)	(0.278)	(0.174)
$ln(Alpha)$	1.312***	1.408***	0.688***
	(0.185)	(0.188)	(0.125)
$PseudoR^2$	0.205	0.122	0.150

*significant at the 0.1 level, ** 0.05, ***0.01.

$N = 254$, Standard errors in parentheses

Table XI: Regression Results: ZIP

	Medical Imaging Patents	Neural Networks Patents	Signal Processing Patents
Papers	0.022***	-0.004	0.006***
	(0.004)	(0.003)	(0.000)
$ln(PSTLabor)$	0.386***	0.789***	0.828***
	(0.055)	(0.050)	(0.018)
Constant	0.469**	-0.893***	0.343***
	(0.199)	(0.198)	(0.065)
<u>INFLATE</u>			
No IEEE Papers	1.074**	0.951**	1.370***
	(0.511)	(0.449)	(0.443)
EE Share of Total Patents	-3.929***	-3.357***	-4.467***
	(1.248)	(1.074)	(1.121)
Ln(Population)	-2.003***	-1.125***	-1.245***
	(0.345)	(0.257)	(0.254)
Constant	28.531***	15.984***	15.441***
	(4.849)	(3.599)	(3.315)
Prob > chi^2	0.000	0.000	0.000

*significant at the 0.1 level, ** 0.05, ***0.01.

$N = 254$, Standard errors in parentheses

Table XII: Regression Results: OLS (including test for anchor)

	Medical Imaging			Neural Networks			Signal Processing		
	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors (WLS)	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors (WLS)	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors (WLS)
Papers	0.405*** (0.100)	0.424*** (0.096)	0.182*** (0.066)	-0.151* (0.077)	-0.125* (0.073)	0.024 (0.060)	1.148*** (0.214)	1.031*** (0.211)	0.449*** (0.133)
Anchor		-1.393 (3.885)	-2.082 (2.999)		5.154*** (1.892)	5.159*** (1.833)		14.649 (13.581)	6.764 (11.142)
Anch*Pap		0.750*** (0.202)	0.828*** (0.213)		0.304*** (0.102)	0.353*** (0.125)		1.841** (0.715)	2.318*** (0.716)
PST Labor	0.053*** (0.012)	0.024* (0.013)	0.057*** (0.012)	0.105*** (0.009)	0.070*** (0.010)	0.051*** (0.011)	0.470*** (0.069)	0.334*** (0.077)	0.483*** (0.067)
Constant	-0.154 (0.463)	0.096 (0.450)	-0.165 (0.234)	-0.057 (0.337)	0.100 (0.327)	0.148 (0.198)	-3.188 (2.370)	-2.058 (2.370)	-1.076 (1.197)
$AdjR^2$	0.406	0.454	0.390	0.542	0.591	0.408	0.647	0.665	0.584

*significant at the 0.1 level, ** 0.05, ***0.01.

$N = 254$, standard errors in parentheses

Table XIII: Regression Results: ZIP (including test for anchor)

	Medical Imaging		Neural Networks		Signal Processing	
	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors	Patents excl. anchors
Papers	0.018***	0.018***	0.001	0.006**	0.007***	0.006***
	(0.004)	(0.005)	(0.004)	(0.003)	(0.000)	(0.001)
Anchor		0.441***		0.953***		0.482***
		(0.103)		(0.119)		(0.043)
$ln(PSTLabor)$	0.604***	0.548***	0.702***	0.446***	0.801***	0.737***
	(0.066)	(0.068)	(0.055)	(0.060)	(0.020)	(0.021)
Constant	-0.614**	-0.464*	-0.812***	-0.215	0.214***	0.349***
	(0.254)	(0.254)	(0.217)	(0.213)	(0.072)	(0.073)
<u>INFLATE</u>						
No IEEE Papers	1.000*	1.015**	0.939**	0.963**	1.378***	1.379***
	(0.529)	(0.528)	(0.454)	(0.447)	(0.454)	(0.449)
EE Share of Total Patents	-4.084***	-4.090***	-3.398***	-3.446***	-4.566***	-4.481***
	(1.365)	(1.363)	(1.093)	(1.084)	(1.158)	(1.133)
Ln(Population)	-1.893***	-1.915***	-1.135***	-1.244***	-1.213***	-1.244***
	(0.357)	(0.357)	(0.258)	(0.255)	(0.257)	(0.255)
Constant	26.975***	27.283***	16.123***	17.640***	15.008***	15.405***
	(5.037)	(5.038)	(3.627)	(3.572)	(3.358)	(3.333)
Prob > chi^2	0.000	0.000	0.000	0.000	0.000	0.000

*significant at the 0.1 level, ** 0.05, ***0.01.

$N = 254$, standard errors in parentheses

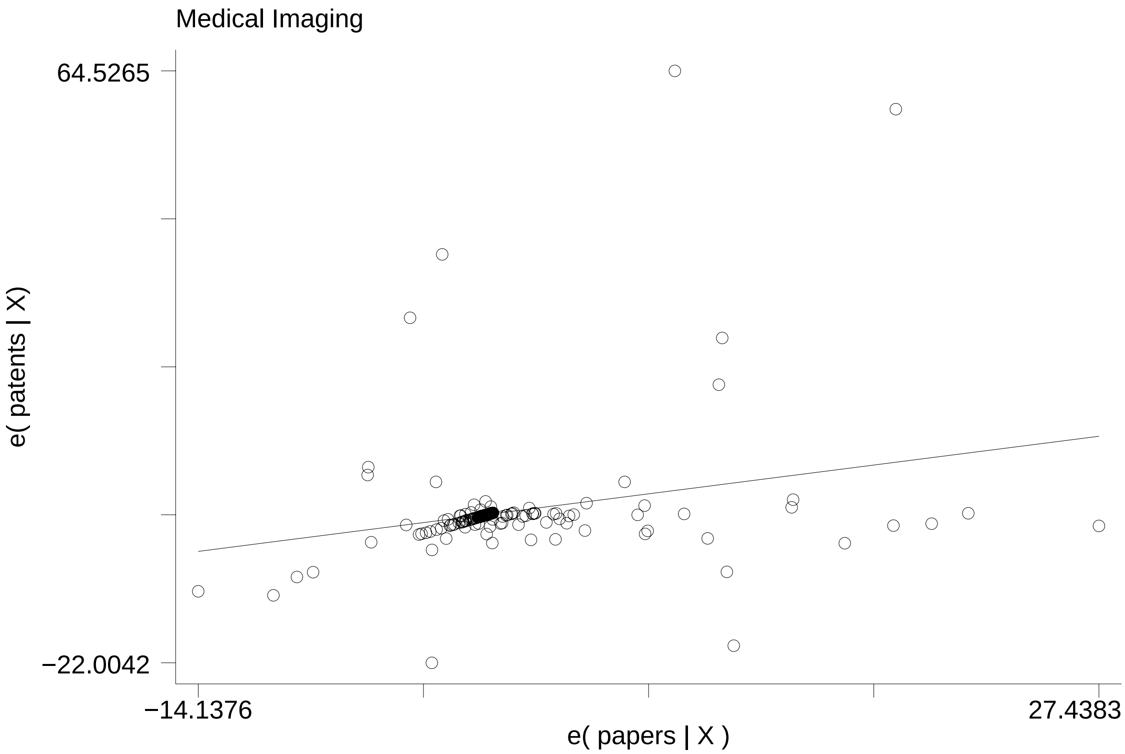

Figure 1: **Medical Imaging**

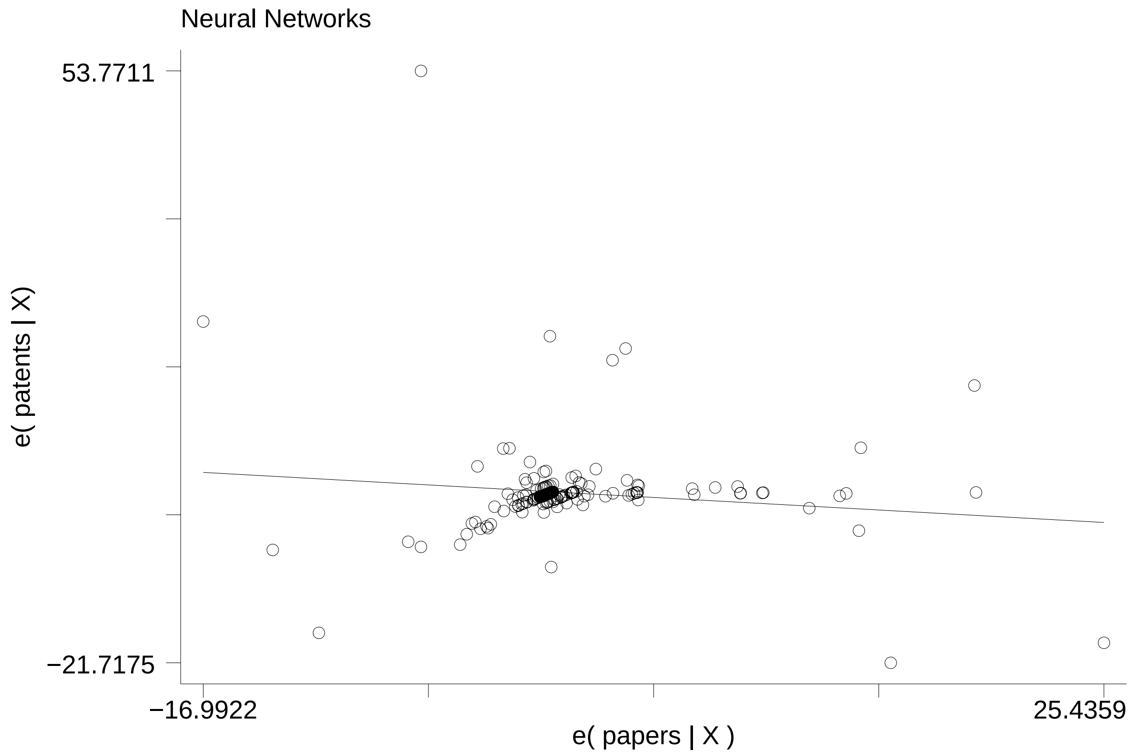

Figure 2: **Neural Networks**

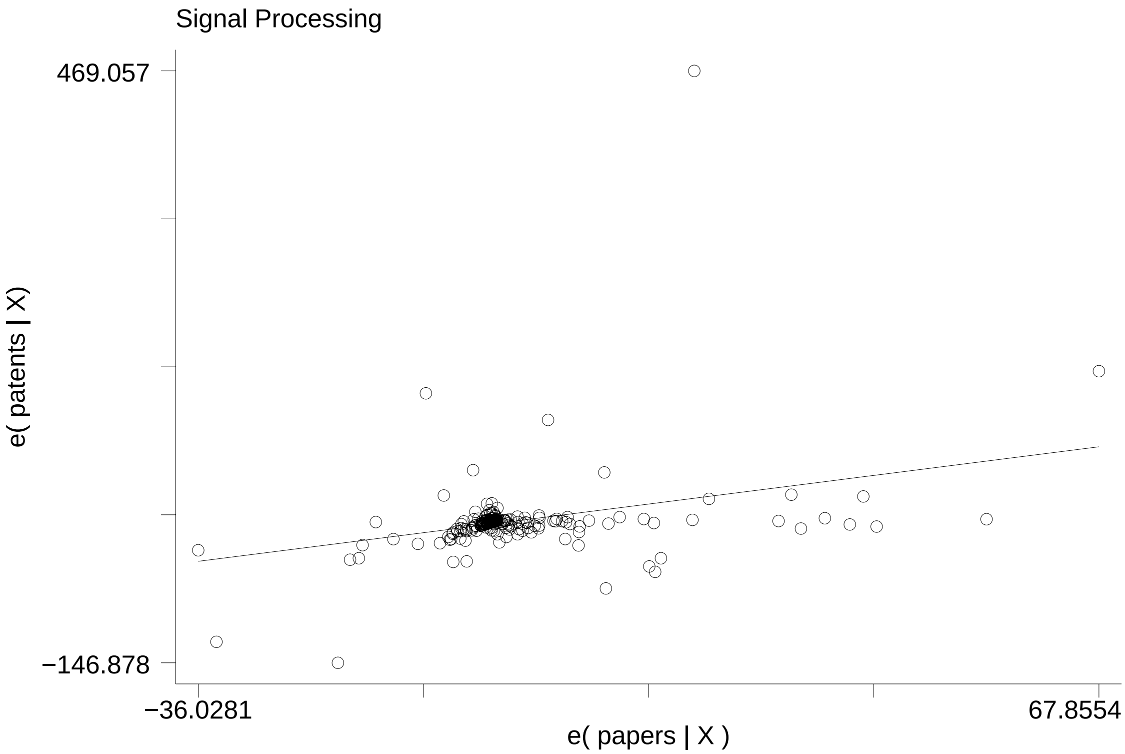

Figure 3: **Signal Processing**