WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 09-12
WHAT EXPLAINS THE QUANTITY AND QUALITY OF LOCAL INVENTIVE ACTIVITY?

Gerald Carlino
Federal Reserve Bank of Philadelphia
and
Robert Hunt
Federal Reserve Bank of Philadelphia

First Draft: November 2008
This Draft: June 2009

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

WORKING PAPER NO 09-12
WHAT EXPLAINS THE QUANTITY AND QUALITY OF LOCAL INVENTIVE ACTIVITY?

Gerald Carlino

Robert Hunt

Federal Reserve Bank of Philadelphia

First Draft: November 2008
This Draft: June 2009

JEL Codes: O31 and R11

Keywords: Agglomeration economies, Knowledge spillovers, Urban density, Innovation, Patents, Academic R&D, Earmarks

This paper was prepared for the 2008 Brookings-Wharton Conference on Urban Affairs.

The authors especially thank Kristy Buzard, Jake Carr, and Bob O'Loughlin for excellent research assistance. Thanks also to Annette Fratantaro for her work on the Compustat variables. Thanks to Colin Webb and Dominique Guellec of the OECD for sharing their patent citations data and answering our many questions. Thanks to Jim Hirabayashi, of the U.S. Patent and Trademark Office, for his gracious responses to our many questions. We also thank Jordan Rappaport, Gilles Duranton and Matt Turner for sharing data with us. This paper has benefited from comments from Edward Coulson, Gilles Duranton, Vernon Henderson, Robert Inman, Janet Pack, Tom Stark, and Bruce Weinberg. We alone are responsible for any remaining errors.

WHAT EXPLAINS THE QUANTITY AND QUALITY OF LOCAL INVENTIVE ACTIVITY?

Gerald Carlino

Robert Hunt

Federal Reserve Bank of Philadelphia

June 2009

Abstract

We geocode a data set of patents and their citation counts, including citations from abroad. This allows us to examine both the *quantity* and *quality* of local inventions. We also refine our data on local academic R&D to explore effects from different fields of science and sources of R&D funding. Finally, we incorporate data on congressional earmarks of funds for academic R&D.

With one important exception, results using citation-weighted patents are similar to those using unweighted patents. For example, estimates of the returns to density (jobs per square mile) are only slightly changed when using citation-weighted patents as the dependent variable. But estimates of returns to city size (urbanization effects) are quite sensitive to the choice of dependent variable.

Local human capital is the most important determinant of per capita rates of patenting. A 1 percent increase in the adult population with a college degree increases the local patenting rate by about 1 percent.

With few exceptions, there is little variation across fields of science in the contribution of academic R&D to patenting rates. The exceptions are computer and life sciences, where the effects are smaller. There is greater variation in the contribution of R&D funded by different sources—academic R&D funded by the federal government generates smaller increases in patenting rates than R&D funded by the university itself. This effect is somewhat stronger for federally funded applied R&D than for basic R&D. We also find small negative effects for cities with greater exposure to academic R&D allocated by congressional earmarks.

We discuss the implications of these results for policy and future research.

1. Introduction

In models of endogenous growth, knowledge, rather than tangible assets, plays a central role in the economic growth of nations. Romer (1990) assumes that economic agents everywhere have free access to the stock of knowledge. Agrawal et al. (2008), among many others, point out that immediate accessibility to knowledge is likely to depend on the geographic proximity of agents. This intuition has been verified in empirical studies of patterns of patent citations (Jaffe, et al. 1993) and in studies of knowledge spillovers among advertising agencies in New York City (Arzaghi and Henderson 2005) and, more generally, in manufacturing (Rosenthal and Strange 2001).

In earlier research, we found additional evidence of such spillovers. In particular we found that the rate of patenting per capita—or *patent intensity*—is about 20 percent higher in a metropolitan area with twice the *employment density* (jobs per square mile) of another metro area (Carlino et al. 2007). In addition to a number of other interesting results, we documented the importance of local research and development (R&D) inputs, in particular human capital, in explaining the inventive productivity of cities.

In this paper, we extend the analysis in a number of important dimensions. First, we introduce a measure of the quality of local inventions—the number of citations a patent receives in patents obtained by subsequent inventors. These "forward" citations have been demonstrated to be correlated with a variety of indicators of value, and they have been used to document the highly skewed distribution in the value of patented inventions. We rely on a relatively new, and underutilized, source of citations—the OECD/EPO Patent Citations Data Set. Using these data, we can determine whether our earlier results are sensitive to these adjustments for the quality of local inventions.

Second, we decompose our data on local academic R&D in a number of important dimensions, including the sources of R&D funding, R&D performed in different fields of science, and the mix of basic vs. applied R&D that is funded. This permits us to test whether the results of academic R&D are indeed homogeneous. Third, we incorporate data on congressional earmarks for R&D that is largely performed by colleges and universities. We are able to compare these earmarks to the overall patterns of federal funding for academic R&D and to test for inefficiencies introduced by the allocation of funds through that process.

Adjusting for the quality of patents does not dramatically change most of the results we found using our simple measure of patents per capita. For example, regardless of whether we use an unweighted or weighted measure of patent intensity, the elasticity associated with employment density is about 0.22. In other words, doubling the employment density of a metropolitan area raises the per capita output of patents by 22 percent.

But some of our results do change. Using unweighted patents per capita, scale (total employment) is not statistically significant in the regressions unless we allow for diminishing returns. With citation-weighted patents per capita, however, the implied elasticity of scale is 0.12 and statistically significant. If we do allow for diminishing returns and we adjust for the quality of inventions, we find that metro areas enjoy increasing returns to scale over a much larger range than we estimated previously. In that case, these returns are exhausted at a population of about 1.8 million. In contrast, when we do not employ citation weighting, our estimates suggest that these returns to scale are exhausted for populations of around 720 thousand.

The presence of an educated workforce is the decisive factor that explains the inventive output of cities, even after controlling for the historical mix of industries and technologies invented. Evaluated at the mean, a 10 percent increase in share of the workforce with at least a college degree raises our measures of patent intensity by about 10 percent. All else equal, a one-standard-deviation increase in our human capital variable is associated with a 30 percent higher patent intensity.

As we found in our earlier research, once we account for local human capital, the effects of incremental increases in local R&D intensities (among private labs, government labs, and academia) are relatively modest. For example, evaluated at the mean, a 10 percent increase in the ratio of private labs to total private establishments raises citation-weighted patent intensity by about 1 percent. A comparable increase in academic R&D intensity has a somewhat smaller effect.

We find very modest variations in the contributions of academic R&D in certain scientific fields. Two exceptions include mathematical and life sciences, which produce fewer patents than other fields (the implied elasticities at the mean are -0.03 and -0.06 respectively). The effect is particularly striking for computer science, especially given the rapid growth in software patenting during this period (Bessen and Hunt 2007). But these results do not necessarily imply fewer inventions; they may simply reflect less reliance on patenting in those fields.

We find that increases in the share of academic R&D funded by the federal government reduce patent intensities relative to R&D funded by the university itself. This effect is relatively large (the elasticity at the mean is about -0.14). In contrast, academic R&D financed by other sources, perhaps including private foundations, is more productive. While it is not clear why federal funds for academic R&D tend to produce fewer patents, it should be remembered that generating patents is not the primary objective of making those funds available.

To further explore this effect, we decomposed the federally funded academic R&D into basic and applied shares. Using these measures, we found a statistically significant negative effect only for the applied portion of federal R&D funding to universities. But the difference in the coefficients on the basic and applied R&D is not statistically significant. While admittedly weak, these results are surprising, given our initial expectation that applied R&D is more closely related to final products and thus more conducive to patenting.

Finally, we find evidence that most academic R&D allocated via congressional earmark (in 1990, at least) did not seem to come at the expense of the major programs that allocate research funds via some form of peer review. Nevertheless, we do find a small negative effect associated with more earmarked academic R&D in metro areas (the elasticity at the mean is about -0.05). There are small variations in this effect depending on which federal agency's budget has been earmarked. Perhaps these funds distract researchers from more promising agendas. Alternatively, it is possible that the projects funded are not ones that rely on the patent system.

The remainder of this paper is organized as follows. Section 2 reviews the literature. Section 3 introduces our data. Section 4 describes the regression strategy and some of our main results. Section 5 presents our new results for academic R&D and for academic earmarks. Section 6 reviews some robustness checks, including tests for unobserved heterogeneity, endogeneity, and spatial dependence. Section 7 reviews a number of policy implications and concludes.

2. Literature

Much of the theoretical literature on urban agglomeration economies has focused on externalities in the production of goods and services rather than invention itself. Nevertheless, the three mechanisms primarily explored in this literature are also relevant for the invention of new goods and services: input sharing, matching, and knowledge spillovers.[1] The first of these points to the sharing of indivisible factors of production, or the benefits of increased variety of differentiated

[1] These themes are developed in the excellent survey by Duranton and Puga (2004). Recent surveys of the empirical literature on agglomeration economies include Eberts and McMillen (1999) and Rosenthal and Strange (2004).

inputs, that occurs in areas with a large number of final-goods producers.[2] A second theory argues that denser urban agglomerations improve learning and the quality of matches among firms and workers.[3] The third strand of theory argues that the geographic concentration of people and jobs in cities facilitates the spread of tacit knowledge. For example, denser locations can hasten learning about new technologies.[4] But there can be too much density in the sense that it may be harder to maintain trade secrets in more dense locations. This potential for poaching may force firms in dense areas to rely on patenting to a greater extent.

In this paper we are agnostic about the precise mechanism by which density affects innovative activity. That is, we do not expect our regressions to be able to sort among the various mechanisms (input sharing, matching, and knowledge spillovers) that likely matter for the innovation process in cities. Our goals are (1) to demonstrate that local job density is empirically relevant in explaining the quality-adjusted volume of local innovative activity, (2) to test for increasing returns to scale (urbanization economies) in cities' output of quality-adjusted innovation and (3) to more precisely quantify the contributions made by the available local R&D inputs, including human capital, and (4) to discuss the implications of these results for public policy.

A full review of the empirical literature on the geographic extent of knowledge spillovers is beyond the scope of this paper, but we will touch on a few especially relevant papers.[5] Several

---

[2] For example, Ciccone and Hall (1996) show how density can give rise to increasing returns in production due to the greater variety of intermediate products available in denser locations. They argue that the positive correlation between employment density and productivity implies that agglomeration economies dominate the congestion effects. See also Helsley and Strange (2002), and Sedgley and Elmslie (2004).

[3] Models of this sort include Glaeser (1999), Wheeler (2001), Helsley and Strange (2004), and Berliant, Reed, and Wang (2006) and its refinement in Hunt (2007). In the latter two models, workers in dense locations are more selective in their matches because the opportunity cost of waiting for a prospective partner is lower. That is because, even though agents are more selective, on average they form matches more quickly. As a result, the average output from matches is higher, and a higher share of the workforce is engaged in productive matches.

[4] See, for example, Glaeser (1999). Some authors argue that local creativity (Florida, 2002) and local entrepreneurship (Acs, 2002) are also conducive to innovation.

[5] See Audretsch and Feldman (2004) for a review of the literature on the geography of knowledge spillovers.

studies have found convincing evidence that spatial proximity mediates knowledge spillovers. Rosenthal and Strange (2001), for example, find that the effects of knowledge spillovers on the agglomeration of manufacturing firms tend to be quite localized, influencing agglomeration only at the ZIP code level.

Looking at innovative activity, Jaffe et al. (1993) and more recently Agrawal et al. (2008) find that nearby inventors are much more likely to cite each other's inventions in their patents, suggesting that knowledge spillovers are indeed localized. In the latter paper, the authors report that every 1,000 mile increase in the distance between two inventors reduces the probability of knowledge flow (as measured by patent citations) by about 2 percent. Arzaghi and Henderson (2005) find that the density of advertising agencies in New York City contributes to information spillovers that enhance productivity, but those spillovers dissipate rapidly with distance among agencies.

Ciccone and Hall (1996) look at the relation between county employment density and productivity at the state level. They find that a doubling of employment density in a county results in an increase in average labor productivity of about 6 percent. More recently, Combes, et al. (2008) revisit the relationship between productivity and density using French data. In addition to using standard historical instruments to deal with endogeneity between productivity and density, they introduce a new set of geological instruments, e.g., depth to bedrock, dominant parent material (the material from which soil forms), topsoil water capacity etc. The idea is that geological characteristics should be fundamental drivers of population settlement patterns. They find that simultaneity bias between density and productivity is relatively small, reducing the impact of density on productivity by about 20 percent. Their preferred estimate for the elasticity of total factor productivity with respect to density is around 3.5 percent.

But why is density important for productivity? Carlino et al. (2007) show that density is important in explaining innovative output, and this may explain the pattern in productivity found in both Ciccone and Hall (1996) and Combes, et al. (2008). Specifically, Carlino et al. (2007) found that, all else equal, a metropolitan area with twice the employment density (jobs per square mile) of another metropolitan area will exhibit a patent intensity (patents per capita) that is 20 percent higher.

Several authors find that patent activity increases with metropolitan area size.[6] But most of these studies do not explain how other city characteristics, such as local density, influence the production of these spillovers. And because these studies do not control for local inputs into the innovation process, such as R&D, or the educational attainment of the labor force, they cannot clearly distinguish between spillovers that are external to individual workers or firms. Carlino et al. (2007) find that after controlling for many local inputs into the R&D process, the benefits of urban scale are realized for cities of moderate size. In fact, with the exception of San Jose, the top 5 percent of metropolitan areas ranked in terms of patent intensity had populations below 1 million.

In addition to returns to scale (urbanization economies), researchers have also investigated whether there are also increasing returns to the size of an industry in a city (localization economies). While we do not touch on that question here, Carlino and Hunt (2007) tested for both urbanization and localization economies in patenting rates in more than a dozen industries. They found evidence of both effects in most industries, and the estimated localization economies were typically comparable to the estimated urbanization economies. This suggests that inter-industry spillovers are often just as important as intra-industry spillovers in explaining local rates of innovation.

A number of studies have looked at the role local inputs into R&D activity play in the

[6] See, for example, Feldman and Audretsch (1999) and O hUallachain (1999).

innovative process. For example, Jaffe (1989), Audretsch and Feldman (1996), and Anselin et al. (1997) found evidence of localized knowledge spillovers from university R&D to commercial innovation by private firms, even after controlling for the location of industrial R&D. Carlino et al. (2007) found that local R&D inputs, especially human capital, contribute to higher patent intensities.

Andersson, Quigley, and Wilhelmsson (2005) find evidence that the expansion of the number of university-based researchers in a local labor market is positively associated with an increase in the number of patents granted in that area. Agrawal and Cockburn (2003) argue that local academic R&D is likely to be more productive, in terms of its contribution to additional patents, in the presence of a large research-intensive firm located nearby—the anchor tenant hypothesis. Taking this effect into account, they report a significant positive correlation between local patents and academic publications in the fields of medical imaging, neural networks, and signal processing.

Economists debate the effects of an area's market structure on the rate of innovation and growth. Chinitz (1961) and Jacobs (1969) argued that the rate of innovations is greater in cities with competitive market structures. Glaeser, et al. (1992) argue that the Marshall-Arrow-Romer (MAR) view implies that local monopoly may foster innovation because firms in such environments have fewer neighbors who imitate them. The empirical literature tends to favor the Chinitz and Jacobs view over the MAR view. Feldman and Audretsch (1999) and Carlino et al. (2007) find that local competition is more conducive to innovative activity than is local monopoly. Glaeser, et al. (1992) find that local competition is more conducive to city growth than is local monopoly.

3. Our Data and Some Descriptive Statistics

As in Carlino et al. (2007), we continue to measure innovations (imperfectly) by using counts of patents obtained by a city's inventors over the years 1990-99. We also employ a rich set of controls for the historical mix of industries and technologies present in a city (see below). These controls help us to address both concerns about the heterogeneity of cities and biases associated with using patents as a proxy for innovation.[7] Since we are primarily interested in explaining variations in inventive productivity, all of our regressions normalize the left-hand-side variable by population. We thus refer to our dependent variable as *patent intensity*.

One concern about using patents as an indicator of innovation is that the value of patents is very highly skewed. Most are not worth very much, while some have values that are higher by several orders of magnitude (see, for example, Harhoff, et al. 1999). Fortunately, there are ways to introduce an adjustment for quality into these counts, just as is done for journal articles: by counting the number of citations a patent receives in subsequent patents. A number of empirical studies document a strong positive correlation between these "forward" citations and the economic value these patents contribute to the firms that own them. For example, Hall, et al. (2005) show that a one-citation increase in the number of patents in a firm's portfolio increases its market value by 3 percent.[8] In addition, these citations present a concrete illustration of knowledge spillovers.[9]

We regress both patent intensity and citation-weighted patent intensity in a metropolitan area on measures of city size, the density of jobs, local market structure, and a set of variables

[7] For a general discussion on the use of patents as indicators, see Griliches (1990)

[8] For further evidence from U.S. patent data, see Trajtenberg (1990). For results correlating survey data to highly cited European patents, see Harhoff et al. (2003) and Gambardella et al. (2008).

[9] In a survey of 1,300 inventors, Jaffe, Trajtenberg, and Fogarty (2000) found that approximately one-half of the patent citations refer to some sort of knowledge spillovers, of which 28 percent correspond to a very substantial spillover. Jaffe et al. (1993) provide evidence that these spillovers are at least initially localized.

that capture the availability of local R&D inputs. To mitigate any bias induced by endogeneity or reverse causation, all the independent variables are lagged—none reflect economic activity after 1990. In a later section of this paper, we investigate these potential biases more closely and find that if any exist, they exert a downward bias on our OLS estimates. In that sense, our OLS estimates can be viewed as conservative. Before presenting the exact specification, we will describe the variables used in our regressions.

The sample consists of 280 metropolitan areas (MAs). Included in this sample are 264 primary metropolitan statistical areas (PMSAs) and nine consolidated metropolitan statistical areas (CMSAs) as they were defined by Census Bureau in 1983 (we employ a lagged definition of metropolitan areas to rule out another potential source of bias in our results). The remaining seven MAs were constructed by aggregating 21 separate PMSAs.[10] We do this because we locate our patents based on a unique match of addresses to a county or city and there is a tendency for a higher number of non-unique matches to occur when cities are nearby.

3.1 The patent data

We assign patents to a metropolitan area using the *residential* address of the first inventor named on the patent.[11] We are able to locate over 581,000 patents granted over the 1990-99 period to inventors living in the U.S. to either a unique county or MA, a match rate of 96 percent. Just over 534,000 (92 percent) of these patents are associated with an urban county.

Figure 1 reproduces the distribution of patents per capita over the 1990s reported in Carlino et al. (2007). In terms of frequencies, the figure implies our data are highly left skewed.[12]

[10] See the appendix to Carlino et al. (2007) for a list of MSAs that were combined. In that paper, we verified that our results did not depend on the inclusion of the ad hoc MAs in the regression sample.

[11] In Carlino et al. (2007), we verified that our results were not sensitive to the choice of the first inventor on the patent.

[12] When analyzing these data in Carlino et al. (2007) we could not reject the null hypothesis of a log normal

The average number of patents per 10,000 of population over the 1990s in our data is about 2. But this varies from as little as 0.07 in McAllen, TX, to as much as 17 in San Jose, a center of the U.S. semiconductor industry.

Our patent citation counts are constructed from the OECD/EPO Patent Citations Database. These are data developed by the European Patent Office (EPO) and the OECD's Patent Statistics Task Force.[13] The underlying data tables span all patent applications published by the EPO, WIPO and Patent Cooperation Treaty (PCT) countries over the years 1978-2001.[14] There are some 6.2 million citations to patent and non-patent prior art in this data set. Over 3.4 million patents have been cited at least once in these data (Webb et al. 2005). U.S. patents account for about a third of this total. We match the U.S. patents in these data to our geo-located patents and extract the total number of citations received by each of these patents.

We rely on this source, rather than the citation data in the NBER Patent Citations Data File (see Hall et al. 2001), because the OECD/EPO data set includes counts of citations to a U.S. patent made in patents obtained in both the U.S. and abroad. In other words, it represents a measure of the worldwide influence of a patent. The NBER data set is an excellent resource, but its counts of patent citations are only those made in other U.S. patent documents; citations in foreign patents are not included. For our analysis, we were concerned that such an omission could be problematic, in particular for industries in which a significant share of R&D is performed outside the U.S. If those industries or technologies are dispersed non-randomly across our MAs (and we believe they are), it is possible that censoring of relevant citations might lead to spurious results.

---

distribution.

[13] We thank Dominique Guellec and Colin Webb for technical assistance with the data. For additional details on these and other OECD patent data, see Webb et al. (2005) and the OECD website for the Directorate for Science, Technology and Industry.

[14] WIPO is the World Intellectual Property Organization.

We are able to match essentially all of our patents to ones contained in the OECD/EPO citation data. Figure 2 plots the distribution of citations in our patent data. As other researchers have found, this distribution is very highly left skewed. About half (49 percent) of patents in our data had not received any forward citations at the time the OECD compiled its data. Almost a quarter (24) of the patents had received only one citation. More than 95 percent of our patents had received five or fewer citations. One-half of 1 percent of our patents have received a dozen or more citations. These are likely the most valuable patents in our data.

3.2 Scale, industrial composition, and density

Our primary measure of city *size* is total civilian employment in 1989 as measured in the data used in the payroll employment survey. These are counts of jobs based on the location and sector of the establishment. In addition to total employment we compute the shares of jobs falling into seven one-digit SIC industry groups, plus federal and local government employment shares.[15]

To investigate the potential effects of local market structure on inventive output, we construct a variable similar to one suggested in Glaeser, et al. (1992)—the number of establishments per worker in the metropolitan area in 1989. These data are derived from the 1989 edition of *County Business Patterns*.[16]

Our measure of job density relies on estimates of the built-up areas of urban counties in 1990. Specifically, we use the land area of the u*rbanized areas* (UAs) contained in our MAs. These are defined as an area with a population of 50,000 or more, comprising at least one place, and the adjacent settled surrounding area with a population density of at least 1,000 per square mile (U.S. Census Bureau, 1994). While UAs are not bounded by county lines, we were able to

[15] We construct these variables from county-level data contained in the1999 vintage of the Bureau of Economic Analysis' Regional Economic Information System (REIS).

[16] This data is based on administrative records at the level of individual establishments. For details, see U.S. Census Bureau (1991) and http://www.census.gov/epcd/cbp/index.html

collect data on urbanized-area land area in nearly every urban county and we used this to construct our measure of the relevant land area in our MAs.

We use this measure of land area because it is a better proxy for the space in which urban labor markets function than is total county land area (Mills and Hamilton 1994). This is especially true for counties in the western U.S. For example, in the 1990 census only 12 percent of the 580,000 square miles of land in MSA counties was categorized as urban in nature. In that year, the urban share of MSA land area varied from less than 1 percent in Yuma, AZ, to 65 percent in Stamford, CT.

But there is a trade-off in this choice of land area, since we use our county-level counts of employment to construct our density measures. To the extent that we are picking up employment occurring outside urbanized areas, our job density measure will be somewhat overstated. In Carlino et al. (2007) we explored this question in detail and conclude that the degree of overstatement is relatively small and is only likely to bias our results downward.[17]

3.3 R&D inputs

We are especially interested in the contribution of local inputs to the rate of innovation in our MAs. In our earlier work we identified four of these inputs: local human capital plus the R&D activity of universities, private firms, and any nearby government laboratories.[18] Because these variables are all highly correlated with the size of cities, we include them in *intensity* form in our regressions. For example, local human capital is measured as the share of the population over 25 years of age with a college degree in 1990. Private R&D is captured with a count of the number

[17] In that paper we ran all of our regressions with an alternative measure of size and density using a residence-based employment variable linked to urbanized areas and found very similar results. We did not rely on such a measure, however, because it significantly understates employment in urbanized areas.

[18] We also controlled for the influence of having many nearby universities, a possible college-town effect, by including in our regressions the ratio of college enrollment to population in the years 1987-89.

of private research labs in 1987 divided by the total number of private establishments.[19] Federal government lab R&D in the years 1987-89 is normalized by the number of federal civilian employees in the MA. R&D performed by academic institutions over the years 1987-89 is normalized by total full-time enrollment at colleges and universities in the MA in those years.

In this paper, we retain three of the four measures as they were constructed in Carlino et al. (2007). But we have reconstructed the academic R&D variable and derived a number of new variables based on those data. The new variables decompose academic R&D into as many as 26 separate fields of study, as defined by the National Science Foundation (2006). We are also able to identify the sources of funding for academic R&D, and we decompose federally funded R&D into basic and applied R&D based on the composition of academic R&D funded by different federal agencies.[20] These new data permit us to explore variations in the effects along each of these dimensions.

3.4 Some facts about academic R&D in our data

The total amount of R&D performed by universities over 1987-89 in our data was $40.5 billion. By comparison, again according to the NSF, private industry funded about $200 billion of R&D over those years. Approximately 600 separate campuses reported a positive amount of R&D in our data. The leading institution, Johns Hopkins University, performed about $1.7 billion in R&D over that period. MIT, the University of Wisconsin, and Stanford each performed over $800 million. The 20 largest performers accounted for about one-third of the total.

One complication in geo-coding academic R&D is that about 20 institutions report their R&D at the system level, but they have active research campuses in more than one location. It is

[19] We geo-coded the location of 1,800 private-sector research labs using the 23rd edition of the *Directory of American Research and Technology*.

[20] These data were constructed from various data sets downloaded from the NSF's WebCaspar server.

not uncommon, for example, for a teaching hospital of a large university to be located well away from its main campus. Collectively these institutions account for about $5 billion in R&D – an eighth of the total. More than 10 of these institutions performed at least $150 million in R&D in 1987-89. In our earlier paper, we allocated R&D for these institutions to specific campuses based on the shares of all advanced degrees awarded at those locations.[21] For this paper, we refined our technique by allocating the R&D in each field of science according to the university's geographic distribution of degrees awarded in that field.

In terms of the sources of funding for academic R&D, the federal government was by far the largest provider of support for academic R&D ($25 billion, 60 percent of the total) in 1987-89. Institutions themselves funded over $7 billion in academic R&D (18 percent). State and local governments contributed over $3 billion (8 percent) and private industry about $2.6 billion (7 percent). Other sources, which we believe include private foundations, accounted for another $2.8 billion (7 percent).[22]

The NSF tabulates the mix of federal funding for academic R&D by various federal agencies in terms of basic or applied R&D. Over the fiscal years 1987-89, slightly more than two-thirds (68 percent) of federal agency R&D funds allocated to universities was categorized as basic. There was considerable variation across agencies. The basic R&D share was highest at the NSF (94 percent) and lowest in Health and Human Services, excluding the National Institutes of Health (39 percent). There was also considerable variation across fields of science: 98 percent of all federal academic R&D funds for astronomy are categorized as basic, while in economics it is less than 40 percent. The share of basic R&D tends to be higher in the physical sciences and

[21] These are the sum of doctorates, master's degrees, and first professional degrees awarded in 1987-89. The latter require at least six years of college work and two years of professional training.

[22] According to the NSF (2006), this category of funds includes grants for R&D from non-profit organizations and voluntary health agencies and gifts from private individuals that are restricted by the donor to the conduct of research, as well as all other sources restricted to research purposes not included in the other categories.

somewhat lower in the life and medical sciences.

We are also interested in the effects of variations in the quality of academic research departments. To do this, we coded the ratings of universities in four fields (engineering, physical sciences, mathematical sciences, and life sciences) from the National Research Council's (NRC) Survey of Scholarly Quality of Faculty for 1982.[23] We took the sum of these ratings over universities in an MA and included this as a control in some of our regressions.

3.5 Congressional earmarks of agency funds for academic R&D

An additional question we wish to explore is the effect, if any, of how federal funding for academic R&D is allocated. In particular, we will test for any differences in inventive productivity that are associated with congressional earmarks of R&D funds.[24] We do this by coding a data set of those earmarks published by the *Chronicle of Higher Education*.[25] Our data span the years 1990-2003. Over those years, according to this source, there were about 11,000 items allocated to one or more academic institutions in appropriations acts. The total amount of funds authorized to be spent over those years was about $17.6 billion.

The aggregate pattern of earmarks is depicted in Figure 3. Academic earmark activity seems to follow the overall pattern of annual appropriations. During the period of binding annual appropriations caps, academic earmarks either fell or remained stable. When those caps expired in the late 1990s, academic earmarks surged. In the final year of our data, academic earmarks had increased five-fold from the 1990 level ($2.1 billion vs. $404 million).[26] This surge in

[23] These data were downloaded from NSF's WebCaspar database in 2004. For additional information and a copy of the survey instrument, see Jones et al. (1982). .

[24] For studies of the political economy of these earmarks, see Savage (1999) and de Figueiredo and Silverman (2006). Savage identifies as the first modern academic earmark a $10 million grant for a veterinary school (from the 1977 appropriations for the Department of Agriculture).

[25] This data was downloaded from the Chronicle's website (http://chronicle.com/stats/pork/) in July of 2007.

[26] Savage (1999) offers an alternative explanation—the collapse of collective attempts by universities to refrain from

earmarks has prompted concerns about the effects of this form of allocation on research output.

Table 2 shows that in 1990 nearly half of the dollar value of academic earmarks (47 percent) was found in the budget for the Department of Agriculture (USDA). This is not surprising given this agency's long history of funding university research. The idea of allocating R&D funds via some form of peer review took hold only shortly after World War II. Prior to that time, most federal research funding was done in the form of agricultural research grants to, for example, support the extension service (Savage 1999).[27]

The other notable concentrations appeared in the appropriations for the Energy and Defense departments and NASA, which accounted for 18, 7, and 4 percent, respectively, of the dollar value of earmarks. In the later years of our data, earmarks in the defense appropriations increased dramatically. As a consequence, academic earmarks on DoD budgets account for nearly a third (31 percent) of all earmarked funds over 1990-2003.[28]

Table 2 also reports the R&D spending of various federal agencies in 1990.[29] If we compare the NSF R&D numbers to the earmarks data, two patterns seem clear. First, the distribution of aggregate federal resources for R&D is very different from the distribution of those resources to academic institutions. For example, compare the shares of federal R&D accounted for by the Department of Defense and the National Institutes of Health (NIH). In terms of aggregate R&D spending, R&D in the DoD budget was 4.5 times larger than that for the NIH. But in terms of R&D allocated to academia, the NIH allocated an amount almost four times

---

seeking academic earmarks around 1993.

[27] For a fascinating description of the institutional details and history of R&D earmarks affecting the USDA, see Law and Tonon (2006).

[28] For a description of the research performed by centers created with two of the largest (and earliest) earmarks of the USDA and DoE budgets, see Mervis (2008).

[29] These are federal R&D obligations for universities & colleges in the 1990 budget, as tabulated by the NSF and downloaded from WebCaspar in the summer of 2008.

greater than that for the DoD.

Second, while only about 4 percent of federal agency R&D support to academia was earmarked in 1990, there was considerable variation across agencies. For example, a majority of academic R&D in the budget for the Department of Agriculture was allocated via earmark. At Energy, 15 percent of academic R&D funding was allocated via earmarks.[30] In contrast, while two-thirds of federal support ($6.1 billion) for academic R&D was funded through the NSF and the NIH in 1990, there were no earmarks in their appropriations in that year. In fact, in all the years of our data, there are no earmarks of NSF funds and only 12 earmarks, for a total of $30 million, of NIH funds (in FY 1996-7).[31]

One conclusion that could be drawn from the table, then, is that most R&D earmarks represent a reallocation of federal funds in *addition to* the vast majority of R&D funds allocated to university researchers based on some form of peer review.[32] The efficiency implications of these earmarks are thus less clear than might first appear. For example, the funded research may simply be directed to questions that are not particularly relevant for private markets. Nevertheless, there might still be an effect on inventive productivity, however, if earmarked funds divert academic manpower from more promising pursuits.

There is some evidence of such effects in terms of research publications. Payne (2002) found that $1 million in R&D earmarks resulted in an increase of 22 publications among the recipient universities but also reduced the average number of citations to articles published by

[30] These shares are simply the ratio of the number for agency earmarks from the *Chronicle of Higher Education* data divided by the R&D funds provided to universities by the federal agency, as reported by the NSF. These percentages should be treated as very approximate since our measures of earmarks and total R&D funding are derived from different sources. Nor is it clear that the NSF counts all of these earmarks as R&D in its data.

[31] But earmarks are not the only means of directing federal research funding, however. See, for example, Payne (2006), who describes the NSF's set aside program, EPSCoR and its effects on research output. And in a detailed analysis of NIH grants Hegde and Mowery (2008) find that about $1.7 billion of $37 billion in research grants to institutions in 2002-2003 were influenced by the geographic composition of the relevant House appropriations subcommittee.

[32] But plentiful earmarks could eventually lead to an unraveling of political support for the major programs that use peer review (for a discussion, see de Figueiredo and Silverman 2007). Our regressions cannot measure such effects.

researchers at those universities. In other words, while quantity increased, quality declined. Hegde and Mowery (2008) find that the geographic composition of the relevant House appropriations subcommittee has some effect on the funding of biomedical research performers (e.g. specific scientists) in the lowest two quartiles of the distribution of research grants funded by the same NIH institute in earlier years. Earlier studies were unable to establish a systematic relationship between a university's success in obtaining academic earmarks and subsequent changes in their academic rankings (de Figueiredo and Silverman 2007).

To test for any effects on local rates of patenting, we include in some of our regressions academic earmarks, again normalized by full-time college enrollment, and similarly normalized values for earmarks from the primary agencies involved.

3.6 Additional control variables

Since we are limited to cross-section regressions, it is extremely important to control for differences across our MAs that are relevant to explaining either potential measurement error or other variations in inventive productivity. As noted earlier, we include in our regressions a set of controls for the historic mix of *industries* (employment shares). In addition, we include a set of controls for the historic mix of *technologies* developed in an MA. We include shares of patents obtained in each MA during 1980-89 that fall into one of six technology groups as defined in Hall et al. (2001).[33] We also include the share of patents obtained in the 1980s by firms in R&D-intensive industries.[34]

In our earlier paper, we considered the possibility that firms that relied more on trade secrets might be less likely to locate in dense cities where greater worker mobility might

[33] The categories are chemicals, computers, medical, electrical, mechanical, and all other. We included shares of the first five in our regressions. We construct these shares using the NBER Patent Citations Data File.

[34] See Carlino et al. (2007) for details on the construction of that variable.

undermine the effectiveness of trade secrets in protecting innovations. In that case, a correlation between patent intensity and job density might reflect *selection* of firms or industries that rely more heavily on patents. To test for this possibility we constructed a variable, based on survey results in Cohen et al. (2000), that captures the relative importance of trade secret protection across different manufacturing industries. We did this by constructing a weighted average of those ratings using the mix of industrial R&D facilities located in our MAs as weights. We include that variable in the analysis here.

We include a number of other control variables. We control for variations in demographics by including the share of the population in 1990 that is of working age. We also include the percent change in employment over the years 1980-89 as a control for the effects of unobserved differences in local economic opportunities on inventive activity. We also include seven dummy variables based on the BEA economic region in which the MA is located (the Rocky Mountain region is omitted).

Table 1 (see appendix) presents the descriptive statistics of the variables used in the analysis to follow. We also examined the correlation coefficients among these variables. The vast majority of the correlations among the variables are well below 0.50. One exception, as should be expected, is MA employment and its square. Despite the high correlation between these variables there is no evidence of a collinearity problem in our regressions—the coefficients on all other variables are not affected by the inclusion or exclusion of the square of MA employment. Given that we believe there is nonlinearity in scale, we prefer a regression that includes both employment size and its square.[35]

---

[35] A table showing these correlations is available from the authors. A Ramsey RESET test (not shown) for omitted variable bias reveals the null hypothesis of no omitted variables is rejected only when we exclude the square of MA employment in our main regressions.

4. Some Results

We begin by summarizing our estimation strategy. We regress the log of either patents or citation-weighted patents per 10,000 of population over the 1990s on the log of MA size and MA density (both measured in terms of jobs), the log of the inverse of average establishment size, the share of the adult population with a college degree, and our measures of local government, academic, and private R&D input *intensities*. In addition, we include controls for the historical mix of local industries and patented technologies, BEA region dummies, and a number of other control variables. All the right-hand-side variables are lagged or beginning-of-period values to minimize the possibility of endogeneity bias.

The main regression results for our two dependent variables are reported in Table 3. The standard errors reported are corrected for potential heteroskedasticity.[36] The results in column 1 can be compared to Tables 2-3 in Carlino et al. (2007). The only other difference in the data used in this research compared to Carlino et al. (2007) is that here we are using an improved measure of academic R&D intensity. There is also one difference in the specification: here we include share of the adult population with a college degree in log form.

4.1 Results for (un-weighted) patents per capita as the dependent variable

Not surprisingly, the coefficients reported here (see Table 4) are very similar to our earlier results. For example, the coefficient on job density, which is the equivalent of an elasticity, is 0.22, as compared to 0.20 in the earlier paper. When we allow for a non-linear relationship (see the top panel of Table 4), the implied optimal job density is in the range of 2,300-2,500 per

[36] A Breusch-Pagan test (not shown) indicates that this is an issue for our regressions. Our regressions are performed using STATA 10. Unless otherwise mentioned we use robust standard errors (White correction).

square mile.[37] These are slightly higher than the estimate (about 2,220) reported in our earlier paper.

Comparing the coefficients on the scale variables, the implied optimal city size is slightly smaller than we found before (473 thousand jobs vs. 500 thousand jobs). Assuming a labor force participation rate of 66 percent, the implied optimal population is about 720 thousand, a rather modest size. Recall that our dependent variable is already normalized by a measure of scale, so this result suggests that there is a range of city size where increasing returns are operating, but those returns are exhausted at a relatively modest scale. Presumably above this city size, diminishing returns to scale have set in.

Relative to our earlier paper, the coefficient on the inverse of average establishment size hardly changes and remains highly significant. The elasticity is quite large—a 10 percent decrease in average establishment size is associated with a 16 percent increase in patent intensity. Thus, for whatever reason, it appears that cities with a more competitive local market structure enjoy higher inventive productivity. Given the data we are using, we cannot distinguish between static (more competition) and dynamic (more new business formation) explanations.

As in our earlier work, local human capital remains the most important explanatory variable by far in our regressions. All else equal, a 10 percent increase in the share of the local adult population with a college degree raises the per capita rate of patenting by 9.7 percent. There is considerable variation in our human capital variable across cities. According to our estimates, a one-standard-deviation increase in the college educated share would increase the patenting rate by 30 percent.

Examining our other measures of local R&D inputs, the coefficients on federal and

[37] To facilitate comparisons with our earlier paper, we include all the variations with respect to scale and density in Table 4. Note that the coefficients and standard errors for all the other variables in those regressions vary only slightly across specifications that include or exclude the square of density or scale

private lab R&D intensities are essentially the same as reported in our earlier paper (Table 3). The implied elasticities (0.01 and 0.10, respectively) are also very similar to our earlier results (see Table 5). The estimated coefficient on academic R&D intensity is lower than we report in Carlino et al. (2007) (0.07 vs. 0.14 in the earlier paper) but the implied elasticities at the mean are essentially the same (0.07 vs. 0.08 before). The key point is that while these effects are quite precisely measured, they suggest relatively modest incremental contributions when compared to the effects of increasing local human capital.[38] This has important implications for policy, as we discuss later.

As with our earlier results, the historical mix of industries and technologies are important in explaining the inventive productivity of cities, at least as measured by patents (Table 5). In some instances the implied elasticities are quite high: a doubling of the manufacturing share of jobs would imply a one-third increase in patenting per capita; a doubling of the historical share of chemical patents would imply a 30 percent increase.

In both of those examples, it is important to exercise care in interpreting the results since manufacturing firms, and chemical producers in particular, are relatively intensive users of the patent system. At least part of those effects reflect the measurement error associated with using patents as an indicator of innovations when there is differential reliance on patents across industries and fields of technology. But it is clear that we can soak up a good deal of this variation using our historical control variables.

4.2 Results for citation-weighted patents per capita

With a few important exceptions, the general pattern of results using our new dependent variable is similar to what we found when using patents per capita as the dependent variable. For

[38] These elasticities are not affected if we exclude the human capital variables from the regressions.

example, there is essentially no difference in the coefficients on employment density (0.22) in columns 1 and 2 of Table 3. But when allowing for diminishing returns, we do find a difference in the implied optimal job density: adjusting for the quality of local inventions reduces our estimate by 200- 400 jobs per square mile (compare columns 2 and 4 in the top and bottom panels of Table 4).

But we do find some important differences in our results. For example, when adjusting for the quality of local inventions, we find a statistically significant linear effect of scale (total employment) on patent intensity. The implied elasticity of scale is about 0.12 (compare columns 1 and 2 in the top and bottom panels of Table 4). That suggests there is at least some evidence of increasing returns to scale.

But as noted earlier, we found evidence of eventually diminishing returns to scale in our regressions on patents per capita. This is also true in the regressions adjusting for patent quality, but diminishing returns appear to set in much further into the tail of the city-size distribution. Taking into account the labor force participation rate, the implied optimal scale in these regressions could be as high as 1.8 million people (compare columns 3 and 4 in the top and bottom panels of Table 4). Here is clear evidence that the results depend in important ways on how inventions are counted.

Differences in the other coefficients are best examined in elasticity form (Table 5). The elasticity of our human capital variable (1.05) is about 10 basis points higher than in our regression using unweighted patents. The elasticity for academic R&D intensity is about the same, but the elasticity on private lab R&D intensity rises about 3 basis points (an increase of 27 percent). Two elasticities in our historical control variables also rise significantly: the manufacturing share of jobs and the share of patents falling into the mechanical category.

We conclude that most of our earlier results are robust to controlling for the quality of inventions. But in some others (e.g., estimates of returns to scale) the estimated effects vary significantly. And, adjusting for the quality of inventions, we find that local human capital and private R&D intensity are somewhat more important than we found earlier.

5. Further Exploration of Academic R&D and Academic Earmarks

In this section we report new results using our additional variables related to the local academic R&D performed on our sample of MAs. In all of these regressions, including additional variables or their interactions has little or no effect on the coefficients or standard errors of the other variables we have already discussed (and continue to include in the regressions). Thus, for the sake of brevity, we will not report those coefficients in the discussion and tables in this section.[39]

5.1 Controlling for the quality of academic institutions

We added to our base specifications our controls for the quality of academic institutions in the MA. Recall that these are counts of quality rankings of academic departments in four areas of science as reported in a 1982 NRC survey of academics in the field. We explored a variety of regressions using these variables. In all of them the estimated coefficients on these quality variables are positive, but they are never statistically significant.

We suspect the reason for this is that the quality of academic departments is already reflected in the academic R&D funding they receive from the federal government and other sources. As evidence, if we exclude our academic R&D intensity variable from the regression (not shown), the estimated elasticity of the aggregated quality measure is 0.04 and statistically significant. A similar result is obtained using our NRC measure of the quality of academic engineering departments. These results could be interpreted as evidence in support of the view

[39] To be explicit all of the regressions in this section include density, scale and the square of those variables. None of the results reported are sensitive to how we include density or scale in the regressions.

that most academic R&D is allocated based on the quality of the researchers who seek funding.

5.2 Differences across fields of science

Our second set of results explores potential variation in the contribution of different fields of science to patenting rates across our sample of cities. We expected to find stronger effects associated with academic R&D in the areas of engineering and in at least a few of the physical sciences (e.g., chemistry). But we found little evidence to support this expectation.

Table 6 reports the important coefficients (in elasticity form) for our R&D intensity variables and a set of variables that reflect the shares of all academic R&D falling into six of seven major categories of science (interdisciplinary and other sciences, as defined by the NSF, are the excluded share). The elasticity of academic R&D intensity is slightly higher than reported in Table 5. The coefficients on most of the fields of science shares, including social sciences, are quite small and are statistically insignificant. The two elasticities that are statistically significant, mathematical sciences and life sciences, take small negative values (-0.02 and -0.06, respectively).

We also ran our regressions using a much finer breakdown of the fields of science into as many as 23 separate categories (Table 7). Including all of these variables has only a slight effect on the coefficients on human capital or the other local R&D intensity variables. But the estimated coefficients on nearly all of the R&D shares were statistically insignificant.[40]

The most robust coefficient was found for the computer science share and it is negative, but quite small. This is an interesting result, given the patterns reported in Bessen and Hunt (2007). They find successful domestic applications for software patents increased at a 16 percent

[40] This does not appear to be due to collinearity in the share variables. The maximum correlation between any two variables was 0.42 (between mechanical and electrical engineering); for most of the others, the correlation was much smaller.

annual rate over the 1990s. But only 2.5 percentage points of this growth could be explained by growth in the employment of computer programmers or engineers and only about 1 percent could be explained by increases in R&D intensity. Nor does it appear that this growth can be explained by the increased academic R&D in computer science.[41]

It is possible that our controls for the share of patents obtained by high-technology firms and the historical mix of patents falling into different categories of technology are masking the effects we expect to see from different categories for academic R&D. We checked this by rerunning the regressions described in table 7, but excluding those control variables. Excluding those controls does change some of the results. For example, if the dependent variable is patents per capita, the elasticities associated with two fields become positive and significant: electrical engineering (0.02) and other life sciences (0.01). In addition to computer science, two others become negative and significant: agricultural sciences (–0.02) and biology (–0.04). The coefficients on the share variables for other fields remain insignificant. If the dependent variable is citation-weighted patents per capita, a statistically significant positive elasticity was found only for electrical engineering (0.03). Negative (and statistically significant) elasticities were found in other engineering (–0.02), computer science (–0.03), agricultural sciences (–0.02), and biology (–0.04).

5.3 Effects by source of R&D funding and basic vs. applied research

Our third set of results documents the variation in inventive productivity, as measured by patenting, of academic R&D based on the source of funding. It is not necessarily clear why the funding source should matter, but we do find evidence that it does. Columns 1 and 3 of Table 8 modify our standard specification by including the shares of academic R&D funded by the

[41] Adjusting for inflation, the annual growth rate in computer science R&D performed by universities over the 1990s was only 3.5 percent.

federal government, state and local governments, industry, and other sources (the omitted share is R&D funded by the university itself). We believe that the other category includes grants from private foundations.

Relative to R&D funded by internal funds, R&D funded by the federal government was less productive in terms of generating patented inventions. The elasticity at the mean (-0.14) is quite large relative to many of the other effects we estimate in our regressions. Second, R&D funded by other sources was modestly more productive relative to internally financed R&D. Evaluated at the mean, the implied elasticities are 0.04 and 0.05, using our two dependent variables. Finally, R&D funded by industry, or state and local governments, appears very slightly less productive, in terms of patenting at the MA level, than R&D funded internally. But those effects are too small to be statistically significant.

It is possible that our results for federal R&D support depend on the kind of R&D that is funded. We usually think of government-funded academic R&D as being more directed at basic rather than applied research.[42] If applied research is closer to final goods markets, we might expect $1 million of applied research to generate more patents than $1 million of basic research. Thus, one possible explanation for this result is that it is an artifact of the federal government funding primarily basic R&D at universities (see section 3.4).

To check this possibility, we decompose federal funding for academic R&D into basic and applied shares based on the variation in those shares across federal agencies and in the distribution of agency R&D funding across universities. This permits us to include in our regressions the basic and applied shares of academic R&D in an MA that are funded by the

[42] The NSF (2006) defines basic research as "systematic study directed toward fuller knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific applications towards processes or products in mind." It defines applied research as "a systematic study to gain knowledge or understanding necessary to determine the means by which a recognized and specific need may be met."

federal government.[43] We include these variables in the regressions reported in columns 2 and 4 of Table 8.

In either specification the coefficients on both variables are negative, but only the coefficient on the applied R&D share is statistically significant. The implied elasticity is small (-0.07). The basic and applied coefficients are quite similar and the difference between them is not statistically significant. While this is admittedly a weak result, it is nevertheless surprising given our assumption that applied R&D is more closely tied to final goods markets, and thus to patents.

We explored one additional possibility—variation in the patent propensity of academic R&D funded by different federal agencies. To the specifications described in columns 1 and 3 of Table 8, we added the shares of federal academic R&D funding of the seven most important funding agencies (Agriculture, Defense, Energy, Health and Human Services, NASA, NIH, and NSF). The only statistically significant effect (not shown) was for the Defense share, with an elasticity of -0.02 at the mean.

5.4 The effects of academic earmarks

Our final set of results explores the potential that academic R&D funded by federal earmarks may have different effects on patent intensity than R&D allocated by other means. As discussed in section 3, it is not obvious that such effects would occur, since it does not appear that, in 1990 at least, these earmarks were diverting resources directly from the primary programs (NIH and NSF) that allocate academic R&D via some form of peer review.[44] To test for such effects, we include in our regressions earmarks in the same form of intensity as we did for academic R&D,

[43] These variables are the product of the basic and applied shares of federally funded academic R&D, at the MA level, constructed from the federal agency data and the share of all academic R&D in an MA that is funded by the federal government.

[44] As mentioned earlier, we recognize that high earmark activity might dilute support for larger NIH and NSF budgets. But our regressions cannot test for such effects.

normalizing by full-time college enrollment in the area. We retain in these regressions our controls for the sources of R&D funding.

The initial results are reported in columns 1 and 3 of Table 9. The estimated elasticities for earmarked funds are negative, but quite small. And when unweighted patents are used, the coefficient is not statistically significant (see column 1). Once we include in the regressions the earmarks associated with the primary agencies experiencing this activity, we find somewhat stronger results (columns 2 and 4 of the table). The elasticity for earmarks in general becomes more negative (about -.05), at least two-thirds the size of the elasticities reported on academic R&D intensity. They are also statistically significant. These effects are consistent with the finding in Payne (2002) that institutions receiving earmarks increased the quantity of their research publications, but the quality of those publications fell.

In a few instances, there is a statistically significant difference in the elasticity of earmarks on particular agency budgets, relative to the effect for all other earmarks. These effects are generally positive and, with the exception of earmarks on the Department of Agriculture budget, are an order of magnitude smaller than the general effect. Thus, any negative effect associated with earmarks on USDA appropriations appears to be about a third smaller than the overall effect of congressional earmarks. All of these results are robust to including our coarse field of science controls in our regressions (not shown).

6. Robustness Checks

Our regressions include a good many control variables, and we have taken considerable care in how we construct them. We have also been careful in how we define our metro areas to minimize issues associated with measurement error or the potential endogeneity of city size. Nevertheless, it is still possible that our techniques have introduced both errors of omission and commission. This section summarizes a number of additional analyses we used to assess the robustness of our results. In particular, we explore an alternative measure of job density, an additional check for potential omitted variable bias, an instrumental variables estimation to address any remaining concerns about potential endogeneity bias, and we explicitly test for spatial dependence.

6.1 The density of knowledge workers

To this point, our measures of employment density reflect the entire workforce of the MA. Very few of those jobs, however, are directly involved in the process of inventing new products or processes. So it is reasonable to ask whether it would be better to focus instead on a measure of occupations consisting of the knowledge workers in an MA. To do that, we re-estimate our regressions using a measure of job density that counts only scientists and engineers living in the urbanized area in 1990.[45] The results are reported in Table 10.

The overall patterns are similar, but there are also many differences in the estimated coefficients. In the first column, using unweighted patents as the dependent variable, the elasticity with respect to S&E density is 0.20, just a bit lower than reported for overall job density in Table 3. But the elasticity is higher (0.26) when we adjust for the quality of inventions. We find some evidence of eventually diminishing returns to S&E density, but the increasing returns would be exhausted at very levels so high that they are not observed in our data set.

Once again we find evidence of eventually diminishing returns to scale. In the first column, the coefficients imply that these returns are exhausted for a city with about 340,000 jobs, or a population of about 515 thousand. This is somewhat smaller than the optimal size reported for un-weighted patents in table 4. The coefficients in the second column (where the dependent variable is citation-weighted patents per capita) imply an optimal city size of about

[45] This variable is constructed from census data. See Carlino et al. (2007) for additional details.

710,000, or a population of about 1.1 million. This is roughly the same size as reported in table 4.

After controlling for scientists and engineers, the implied elasticities on local human capital are a bit lower than we report in Table 5. In the regression using unweighted patents, the elasticity at the mean is about 0.89; using citation-weighted patents, it is about 0.95. The elasticities on private lab intensities fall about 2 basis points. The elasticity associated with average establishment size is also a bit smaller than reported in table 3: it is 1.46 in the regression using unweighted patents and 1.29 in regressions using citation-weighted patents.

6.2 Omitted variable bias

Given the many control variables included in our regressions, we are not particularly concerned about this form of bias.[46] But one way of soaking up any potential remaining bias of this sort is to include in our regressions a lagged value of the dependent variable. For the unweighted dependent variable we include (in logs) the average rate of patenting per 10,000 of population over the years 1985-89. For our quality-adjusted dependent variable we include (again in logs) the average rate of citation-weighted patents granted over the years 1980-89. Results are reported in Table 11.

Including lagged values of the dependent variable has the expected effect of reducing the coefficients on both our scale and density variables. In the unweighted patent regressions, the coefficient on one of our scale variables is no longer statistically significant. The coefficients on our density variables are reduced by 4-8 basis points. Our other variables of interest continue to be statistically significant, although the size of the coefficients is typically lower. In general, there appears to be less of an effect on the coefficients when we use citations to adjust for the

[46] As noted earlier, a Ramsey RESET test on our main specifications does not reject the null of no omitted variables.

quality of patents. We conclude, based on this harsh test that potential omitted variable bias does not seem to explain our results.

6.3 Potential endogeneity

We now consider the possibility that the MA employment, MA employment density, and the share of the population with a college degree are endogenous variables. Recall that our main specification relies on lagged values of the independent variables, and even lagged (and fixed) definitions of metropolitan areas, to minimize this possibility. Nevertheless, the concern remains.

As noted by Combes, et al. (2008), since Ciccone and Hall (1996) at least, it is standard practice to use long lags of population density (and population scale) as instruments for employment density (and employment size). We also use lags of the share of the population with a college degree as an instrument for the share of the MA population with a college degree. A good instrument must be correlated with the endogenous explanatory variable (instrument relevance), and the instruments must be contemporaneously uncorrelated with the residuals (instrument exogeneity).

As pointed out by Combes, et al. (2008), it’s highly likely that the spatial distribution of population, employment, and college share will exhibit persistence through time (instrument relevance), but that local determinants of current innovative activity will differ from those of the distant past (instrument exogeneity). The square of the log of MA job density in 1970 and a dummy variable for the significant presence of hills and mountains are used as the instruments for MA employment density. Similarly, the 1940 population of the MA (in logs), and miles of planned highways in 1947 (in logs), and the square of those terms are used as instruments for MA employment size (and the square of employment size). The share of the 1940 population with a college degree is used as an instrument for human capital.

Table 12a shows the results of a variety of regressions when we instrument for MA employment density. Column (1) in the table reports the results of an IV estimation of the log of patents per capita on the log on employment density when no other endogenous or exogenous regressors (except for regional fixed effects) are included in the regression. Column (2) in the table shows the findings when the other exogenous variables are added to the IV regression, whereas column (3) gives the results when the other (potentially) endogenous variables are added to the IV estimation as well. The last three columns of Table 12a are identical to the first three columns of the table, except the log of citation-weighted patents per capita replaces the log of patents per capita as the dependent variable.

As Table 12a shows, the first-stage *F* statistics for our potentially endogenous regressors are well above the rule of thumb (*F*-statistic of at least 10) for strong instruments suggested by Staiger and Stock (1997).[47] Table 12a reports the results on the log of employment density from the second-stage instrumental variables (IV) regressions. Note that we lose more than 50 observations because of the lack of historical data for our instruments.

The estimated coefficients on our employment density variable are somewhat larger than those reported in Table 3 and remain statistically significant. Of course, if our instruments are also endogenous our parameter estimates may still be biased. Since we have more instrumental variables than endogenous, regressors, we use the Hansen *J*-test to verify that the instruments are uncorrelated with the error term. The *p* values from those tests are well above 0.30. The endogeneity tests do not reject the null hypothesis of equality of our instrumented coefficients in the OLS and IV regressions (the *p* values are 0.15 or greater in every instance), indicating that OLS estimations are unbiased and more efficient than the IV estimations.

[47] These statistics also exceed the critical values of the "size" test for weak instruments and the small sample bias test for instrumental variables regressions in Stock and Yogo (2005). To conserve on space, we do not present the results of the first-stage regressions or the second-stage results for any variable other than the variable being instrumented for.

The results reported in Table 12b are similar to those reported in Table 12a, except that we now instrument for MA employment size (and its square). The first-stage *F* statistics indicate that our instruments are strong. The estimated coefficients on our employment size variable are somewhat larger than those reported in Table 3 and are statistically significant in only two of the six specifications.[48] The coefficients on the square of MA employment variable are significant in three of six specifications. This is not surprising given that the IV estimates are typically less efficient than OLS. For the most part, the tests for endogeneity bias for this variable do not indicate that the OLS coefficients are biased. Only in the most parsimonious specification is the null rejected, and in that case, the results suggest our OLS coefficients are biased downward.

Finally, the results reported in Table 12c are similar to those reported in Table 12a, except that we now instrument for human capital. The first-stage *F* statistics indicate that our instruments are strong. The estimated coefficient on the log of the share of the population with a college degree variable is positive and significance in all cases. Since the regressions summarized in the table have a single instrument and a single included endogenous regressor, the coefficients reported in Table 12c are exactly identified, and the Hansen *J*-test is not applicable.[49]

The endogeneity tests reject the null hypothesis of equality of our instrumented coefficients in the OLS and IV regressions (the *p* values are 0.08 or smaller in every instance), indicating that the OLS estimations are biased. However, a comparison of the results for the human capital variable in the OLS regression (an elasticity of about unity) with those reported in Table 12c (an elasticity of at least 2) indicates that to the extent that there is endogeneity bias, the

[48] But the p values (not shown) for our scale variables in the specifications in columns 1 and 4 of the table, are 0.12 and 0.13 respectively.

[49] We also conducted over-identification tests in IV regressions by including a second instrument (log of inches of annual rain fall) and we never rejected the null hypothesis of exogenous instruments. However, the *F*-statistics in the first-stage regression including the other exogenous variables are about 9.0. The coefficients are qualitatively similar to the ones reported in Table 12c. Endogeneity tests consistently reject the null hypothesis of exogeneity for the human capital variable.

bias works against the maintained hypothesis that high-skilled workers positively affect local innovative activity.

6.4 Spatial dependence

There is a very high degree of spatial inequality in the distribution of patent activity. Patenting tends to be highly concentrated in the metropolitan areas of the Northeast corridor, around the Research Triangle in North Carolina, and in California's Silicon Valley. Even though the coefficients on our regional dummy variables are typically insignificant, this clustering of innovative activity suggests that there could be strong spatial dependence at a more localized level and, if so, it should be controlled for in our empirical analysis.

The conjecture, then, is that patent intensity in one MA may be highly correlated with patent intensity in nearby MAs. The consequences of spatial autocorrelation are the same as those associated with serial correlation and heteroskedasticity: When the error terms across MAs in our sample are correlated, the OLS estimation is unbiased but inefficient. However, if the spatial correlation is due to the direct influence of neighboring MAs, the OLS estimation is biased and inefficient (Anselin, 1990). The literature suggests two approaches to dealing with spatial dependence. In the first approach, spatial dependence is modeled as a spatial autoregressive process in the error term. The second approach models the spatial dependence in patenting activity via a spatially "lagged" dependent variable.

Following Anselin and Hudak (1992), we perform three tests for spatial autocorrelated errors: Moran's I test, the Lagrange multiplier (LM) test, and a robust Lagrange multiplier test (robust LM). We also perform two tests for the spatial lag model (LM test and a robust LM test). The Moran's I test is normally distributed, while the LM tests are distributed χ^2 with *k* and one degree of freedom, respectively.

We estimate each of the specifications previously reported in Table 3 using these various tests for spatial dependence. The results are summarized in Table 13. The null hypothesis of either a zero spatial error or zero spatial lag cannot be rejected in any specification (Moran's I test, the LM, or the robust LM test). Thus, spatial dependence does not appear to be a concern in our regressions.

7. Conclusions and Policy Implications

A number of potentially important policy implications seem to follow naturally from our results. What, if anything, should local policymakers do to stimulate local innovative activity? The answer depends, in part, on who benefits from innovative activity. A metropolitan area might be highly innovative, but if the benefits of this innovation largely occur in other regions, local policymakers might have too little incentive to support local innovative activity. That would suggest a role for making policy at a national rather than local level. But this begs the question—what policy instruments are important and who should decide how they should be used?

While this paper does not consider the extent to which patenting stimulates local growth, results found in Carlino and Saiz (2008) are instructive. While measuring the effect of patenting on local growth was not the main purpose of that paper, the results do suggest that more patents obtained by local inventors are associated with more local job growth. Using the estimates reported in Table 5 of that paper, if a city could double its rate of patenting per capita, the increase in local employment over a decade would be 1.9 percentage points higher. This is certainly an economically significant effect, but in itself, it does not tell us about the magnitude of any growth induced in other areas.

Returning to the results in this paper, the most significant policy levers policymakers at any level of government should consider are ones that influence the accumulation of human capital. It is by far the most important variable in explaining the overall rate of inventive activity in cities, even after controlling for other R&D inputs and other city characteristics. The estimated marginal effects of increasing local human capital are nearly an order of magnitude larger than almost any other variable in our analysis.

A second important finding (also reported in our earlier work) is an extremely robust inverse relationship between a city's patenting rate and the average size of its business establishments. The implied elasticities are again quite large. Unfortunately, the limitations of our data preclude us from speculating on the exact channel that explains this relationship. Is it the static benefits of more competitive local labor markets as suggested by Chinitz (1961) and Jacobs (1969)? Or are we identifying the effects of new business formation stimulated by drastic innovations? This is a very important question for future research and for policy design.

Third, we demonstrate that city size and job density are both empirically relevant in explaining the inventive productivity of cities. The marginal effects of density are of about the same magnitude as those for scale, or even larger. While these results are consistent with theories about matching externalities in labor markets (Berliant et al. 2006, Hunt 2007), more theoretical and empirical work is required before precise policy implication can be suggested. And to identify the exact relationships at work, empirical studies must almost certainly be done using panel data and quite likely at the level of individual matches between workers and firms.

Nevertheless, labor markets do seem a very sensible focus for innovation policy. It is not uncommon for researchers to point to characteristics of U.S. labor markets to explain why U.S. labor productivity growth has outpaced that of many other developed countries over the last 15 years. Part of those gains in productivity are almost certainly attributable to the development of new products and services—in other words to innovation. Yet discussions about innovation policy are typically dominated by questions of how to allocate scarce public resources to particular regions in order to stimulate R&D or foster “clusters.” Such policies may prove complementary to, but are certainly not substitutes for, policies that encourage people to obtain education and for labor markets to put those skills to their best use.

Fourth, we find evidence of increasing returns to scale in the local rate of invention, but we also find that these returns are eventually exhausted. In our earlier work, which relied on unweighted patent counts, we found that these returns were exhausted at a relatively modest city size, roughly at the mean of cities in our data. In this paper, we find that adjusting for the quality of inventions using citation weights more than doubles our estimate of the city size at which increasing returns are exhausted. We suspect, however, that more precise and robust results can be obtained by working with panel rather than cross-section data. This is an important topic for future research.

Fifth, we continue to find that increases in local academic R&D generate a modest incremental contribution to the local rate of invention. With only one robust exception—computer science—we found remarkably little variation in the contribution of academic R&D in different scientific fields to local patenting. An increase in academic R&D in that field generates a relatively small increase in the patenting rate. This is likely a historical artifact of the field's lack of reliance on patents, but it stands in contrast to the very rapid growth in software patenting that occurred in the 1990s.

Sixth, we find that there are significant variations in the marginal contribution of academic R&D to patenting rates depending on the source of its funding. Interestingly, we find that “other” sources are the most productive, while federally funded R&D is the least productive, as measured by changes in the local invention rate. Of course we expect that government-sponsored R&D is more basic in character, perhaps several more steps removed from the creation of new products or processes. But our more disaggregated results show that applied academic R&D funded by the federal government was no more productive than the basic R&D and might even be less productive. Clearly there is room for further research on this question.

Finally, we are able to identify a modest negative effect associated with congressional earmarks of federal funds for academic R&D. To our knowledge, this is the first finding of an effect of these earmarks on research productivity as measured by the local rate of invention. This result is somewhat surprising since we could verify that the most important sources of academic R&D allocated via peer review (NIH and NSF) were largely immune to these earmarks during the period we studied. In other words, it appears that, for the most part, the earmarks represented a net addition of R&D funds, rather than a re-allocation of funds. Nevertheless, our findings would be consistent with a less direct channel: the possibility that researchers are distracted from more productive activities. That would be consistent with the results for the quantity and quality of research publications reported in Payne (2002).

REFERENCES

Acs, Zoltan 2002. *Innovation and the Growth of Cities*, Cheltenham: Edward Elgar Publishing.

Agrawal, Ajay and Iain Cockburn. 2003. "The Anchor Tenant Hypothesis: Exploring the Role of Large, Local, R&D-Intensive Firms in Regional Innovation Systems," *International Journal of Industrial Organization,* Vol. 21 (9), pp. 1217-53.

Agrawal, Ajay, Devesh Kapur, and John McHale. 2008. "How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data," *Journal of Urban Economics*, Vol. 64 (2), pp. 258-69.

Andersson, Roland, John M. Quigley, and Mats Wilhelmsson. 2005. "Higher Education, Localization and Innovation: Evidence from a Natural Experiment," CESIS Working Paper No. 26.

Anselin, Luc, Attila Varga, and Zoltan Acs. 1997. "Local Geographic Spillovers Between University and High Technology Innovations," *Journal of Urban Economics*, Vol. 42 (3), pp. 422-48.

Anselin, Luc, and Sheri Hudak. 1992. "Spatial Econometrics in Practice: A Review of Software Options," *Regional Science and Urban Economics,* Vol. 22 (3), pp. 509-36.

Anselin, Luc. 1990. "Some Robust Approaches to Testing and Estimation in Spatial Econometrics," *Regional Science and Urban Economics,* Vol. 20 (2), pp. 141-63.

Arzaghi, Mohammad, and J. Vernon Henderson. 2005. "Networking off Madison Avenue," mimeo, Brown University.

Audretsch, David B., and Maryann P. Feldman. 2004. "Knowledge Spillovers and the Geography of Innovation," in J.V. Henderson and J-F. Thisse, eds. *Handbook of Regional and Urban Economics,* Vol. 4 (Cities and Geography). Amsterdam: North-Holland, pp. 2713-39.

Audretsch, David B., and Maryann P. Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, Vol. 86 (3), pp. 630-40.

Berliant, Marcus, Robert R. Reed, III and Ping Wang. 2006. "Knowledge Exchange, Matching, and Agglomeration," *Journal of Urban Economics,* Vol. 60 (1), pp. 69-95.

Bessen, James and Robert M. Hunt. 2007. "An Empirical Look at Software Patents," *Journal of Economics and Management Strategy,* Vol. 16 (1), pp. 157-89.

Carlino, Gerald A., Satyajit Chatterjee, and Robert M. Hunt. 2007. "Urban Density and the Rate of Invention," *Journal of Urban Economics,* Vol. 61 (3), pp. 389-419.

Carlino, Gerald A. and Robert M. Hunt. 2007. "Innovation Across U.S. Industries: The Effects of Local Economic Characteristics," Federal Reserve Bank of Philadelphia Working Paper No. 07-28.

Carlino, Gerald A., and Albert Saiz. 2008. "City Beautiful," Federal Reserve Bank of Philadelphia Working Paper No. 08-22.

Ciccone, Antonio, and Robert E. Hall. 1996. "Productivity and the Density of Economic Activity," *American Economic Review,* Vol. 86 (1), pp. 54-70.

Chinitz, Benjamin. 1961. "Contrasts in Agglomeration: New York and Pittsburgh," *American Economic Review,* Vol. 51 (2), Papers and Proceedings of the Seventy-Third Annual Meeting of the American Economic Association, pp. 279-89.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. 2000. "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (OR NOT)," NBER Working Paper No. 7552.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2008. "Estimating Agglomeration Economies with History, Geology, and Worker Effects," CEPR Discussion Paper No. 6728.

de Figueiredo, John M. and Brian S. Silverman. 2006. "Academic Earmarks and the Returns to Lobbying," *Journal of Law and Economics*, Vol. 49 (2), pp. 597-625.

de Figueiredo, John M. and Brian S. Silverman. 2007. "How Does The Government (Want To) Fund Science? Politics, Lobbying And Academic Earmarks," NBER Working Paper No. 13459.

Directory of American Research and Technology, 23rd Edition. 1988. New York: R.R. Bowker.

Duranton, Giles and Diego Puga. 2004. "Micro-Foundations of Urban Agglomeration Economies," in J. V. Henderson and J-F. Thisse, eds., *Handbook of Regional and Urban Economics,* Vol. 4. (Cities and Geography) Amsterdam: North-Holland, pp. 2063-2117

Eberts, Randall W., and Daniel P. McMillen. 1999. "Agglomeration Economies and Urban Public Infrastructure," in P. Cheshire and E. Mills, eds., *Handbook of Regional and Urban Economics,* Vol. 3. (Applied Urban Economics). New York: Elsevier Science, pp. 1455-95.

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

Florida, Richard. 2002. *The Rise of the Creative Class:* And How It's Transforming Work, Leisure, Community and Everyday Life, New York: Basic Books.

Gambardella, Alfonso, Dietmar Harhoff, and Bart Verspagen. 2008. "The Value of European Patents," CEPR Discussion Paper No. 6848.

Glaeser, Edward L. 1999. "Learning in Cities," *Journal of Urban Economics*, Vol. 46 (2), pp. 254-77.

Glaeser, Edward L., Heidi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities," *Journal of Political Economy*, Vol. 100 (6), pp. 1126-52.

Griliches, Zvi. 1990. "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, Vol. 28 (4), pp. 1661-1707.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," NBER Working Paper No. 8498.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg, 2005. "Market Value and Patent Citations," *RAND Journal of Economics,* Vol. 36 (1), pp. 16-38.

Harhoff, Dietmar, Francis Narin, F. M. Scherer, and Katrin Vopel. 1999. "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, Vol. 81 (3), pp. 511–15.

Harhoff, Dietmar, F. M. Scherer, and Katrin Vopel. 2003. "Citations, Family Size, Opposition and the Value of Patent Rights," *Research Policy*, Vol. 32 (8), pp. 1343-63.

Hegde, Deepak and David C. Mowery. 2008. "Politics and Funding in the U.S. Public Biomedical R&D System," *Science*, Vol. 322 No. 5909, pp. 1797-8.

Helsley, Robert W. and William C. Strange. 2004. "Knowledge Barter in Cities," *Journal of Urban Economics*, Vol. 56 (2), pp. 327-45.

Helsley, Robert W. and William C. Strange. 2002. "Innovation and Input Sharing," *Journal of Urban Economics,* Vol. 51 (1), pp. 25-45.

Hunt, Robert M. 2007. "Matching Externalities and Inventive Productivity," Federal Reserve Bank of Philadelphia Working Paper No. 07-7.

Jacobs, Jane. 1969. *The Economy of Cities*. New York: Vintage Books.

Jaffe, Adam B. 1989. "Real Effects of Academic Research," *American Economic Review,* Vol. 79 (5), pp. 957-70.

Jaffe, Adam B., Manual Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, Vol. 108 (3), pp. 577-98.

Jaffe, Adam B., Manual Trajtenberg, and Michael S. Fogarty. 2000. "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors," *American Economic Review*, Vol. 90 (2), Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association, pp. 215-18.

Jones, Lyle V., Gardner Lindzey, and Porter E. Coggeshall. 1982. *An Assessment of Research-Doctorate Programs in the United States: Engineering*. Washington, D.C.: National Academies Press.

Law, Marc T. and Joseph M. Tonon. 2006. "The Strange Budgetary Politics of Agricultural Research Earmarks," *Public Budgeting & Finance*, Vol. 26 (3), pp. 1-21.

Mervis, Jeffrey. 2008. "Building a Scientific Legacy on a Controversial Foundation," *Science*, Vol. 321 No. 5888, pp. 480-3.

Mills, Edward S. and Bruce W. Hamilton. 1994. *Urban Economics*. New York: Harper Collins College Publishers.

National Science Foundation. 2006. *Federal Funds for Research and Development: Fiscal Years 2003–05, NSF 06-313*. Arlington, VA.: National Science Foundation.

O hUallachain, Breandan. 1999. "Patent Places: Size Matters," *Journal of Regional Science*, Vol. 39 (4), pp. 613-36.

Payne, Abigail A. 2002. Do US Congressional Earmarks Increase Research Output at Universities?" *Science and Public Policy*, Vol. 29 No. 5, pp. 314-30.

Payne, Abigail A. 2006. "Earmarks and EPSCoR: Shaping the Distribution, Quality, and Quantity of University Research," in David Guston and Daniel Sarewitz, eds., *Shaping Science and Technology Policy: The Next Generation of Research*. Madison, WI: University of Wisconsin Press.

Romer, Paul M. 1990. “Endogenous Technological Change, “*Journal of Political Economy*, Vol. 98 (5), pp. S71-102.

Rosenthal, Stuart S. and William C. Strange. 2004. “Evidence on the Nature and Sources of Agglomeration Economies,” in J. V. Henderson and J-F. Thisse, eds., *Handbook of Urban and Regional Economics*, Vol. 4. (Cities and Geography) Amsterdam: North-Holland, pp. 2119-71.

Rosenthal, Stuart S. and William C. Strange. 2001. “The Determinants of Agglomeration,” *Journal of Urban Economics*, Vol. 50 (2), pp. 191-229.

Savage, James D. 1999. *Funding Science in America: Congress, Universities, and the Politics of Academic Pork Barrel*. Cambridge: Cambridge University Press.

Sedgley, Norman and Bruce Elmslie. 2004. “The Geographic Concentration of Knowledge: Scale, Agglomeration, and Congestion in Innovation across U.S. States,” *International Regional Science Review*, Vol. 27 (2), pp. 111-37.

Staiger, Douglas and James H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments,” *Econometrica*, Vol. 65 (3), pp. 557-86.

Stock James H. and Motohiro Yogo. 2005. “Testing for Weak Instruments in Linear IV Regression,” in J.H. Stock and D.W.K. Andrews, eds., *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg.* Cambridge: Cambridge University Press, pp. 80–108.

Trajtenberg, Manuel. 1990. “A Penny for Your Quotes: Patent Citations and the Value of Innovations,” *The Rand Journal of Economics*, Vol. 21 (1), pp. 172-87.

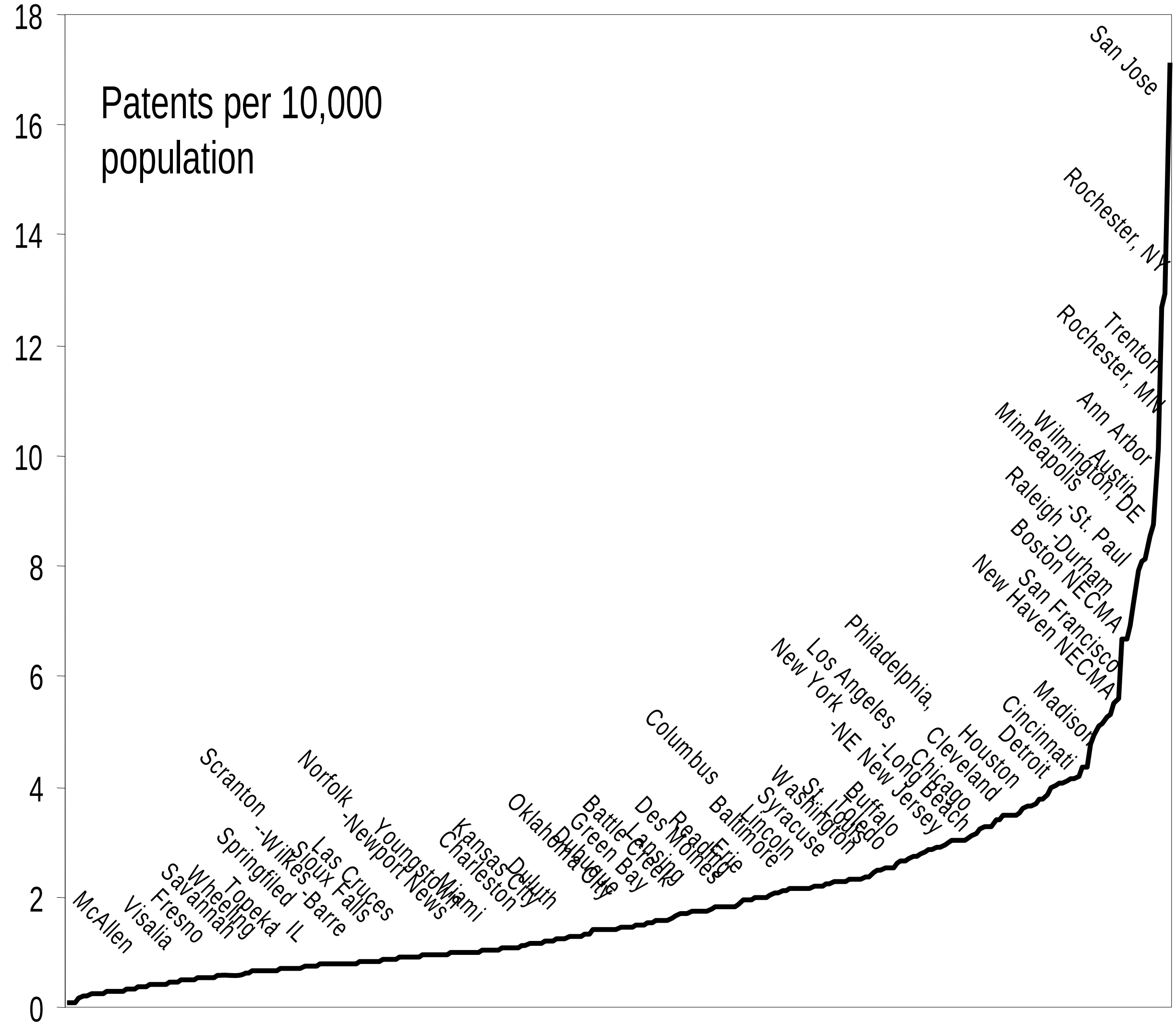
U.S. Census Bureau. 1991. County Business Patterns, 1989. Washington, D.C.: U.S. Government Printing Office.

U.S. Census Bureau. 1994. *Geographic Areas Reference Manual*. http://www.census.gov/geo/www/garm.html

Webb, Colin, Hélène Dernis, Dietmar Harhoff, and Karin Hoisl. 2005. “Analyzing European and International Patent Citations: A Set of EPO Patent Database Building Blocks,” OECD STI Working Paper No. 2005/9.

Wheeler, Christopher H. 2001. "Search, Sorting, and Urban Agglomeration" *Journal of Labor Economics*, Vol. 19 (4), pp. 879-99.

Figure 1: Patent Intensity Across MAs



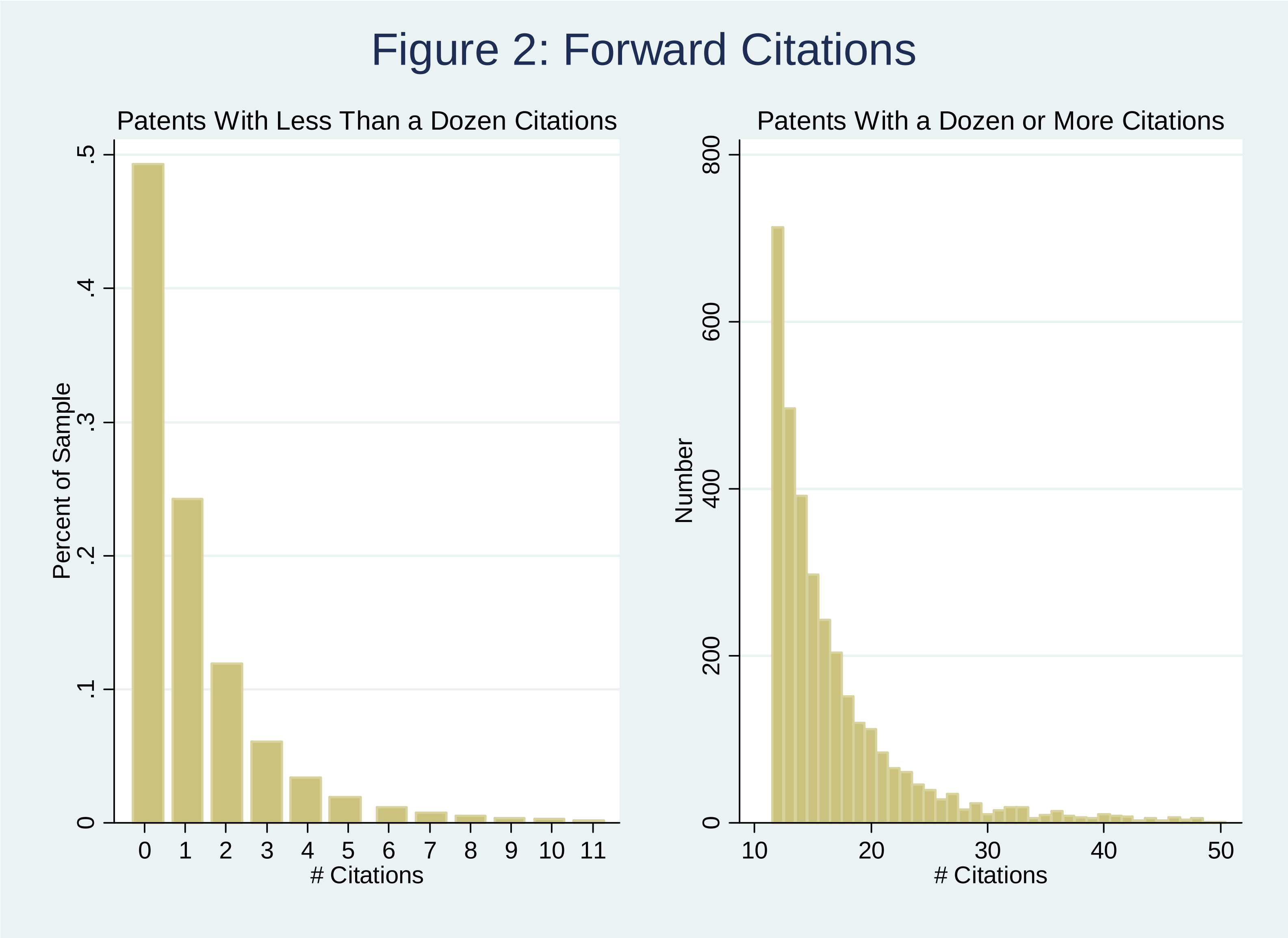
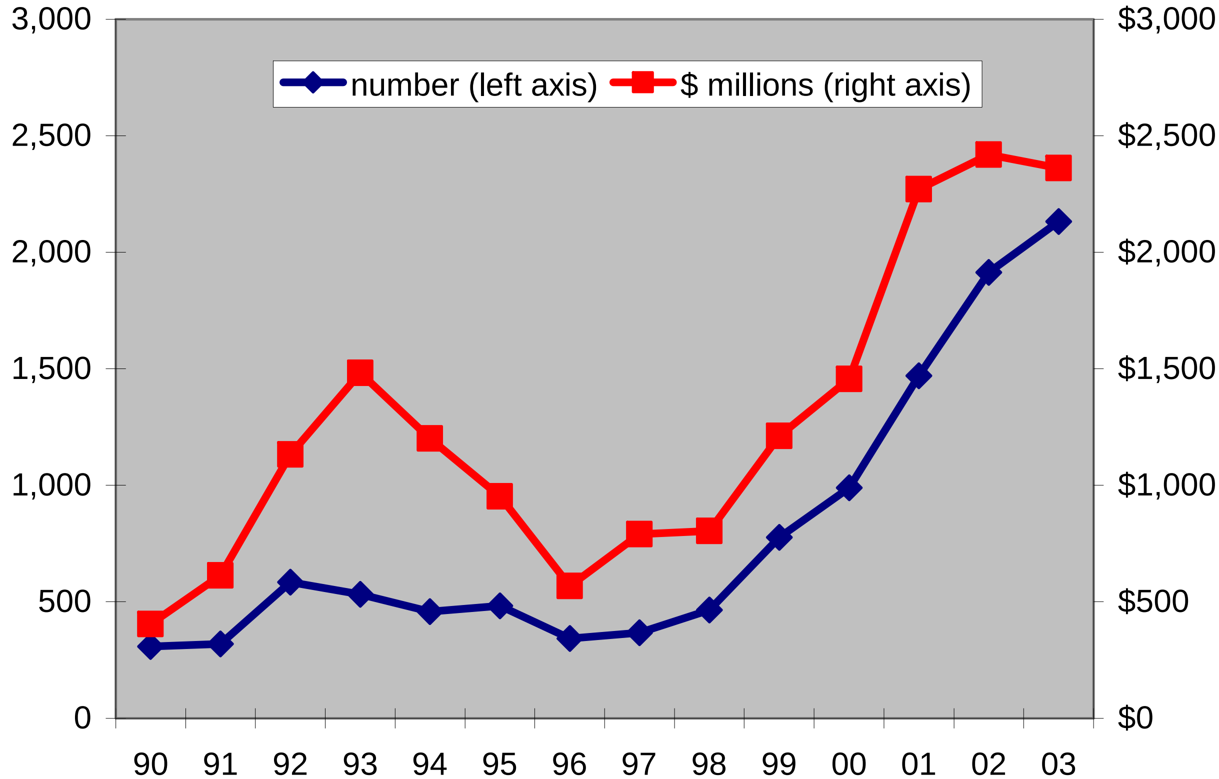


Figure 2: Forward Citations
Patents With Less Than a Dozen Citations
Percent of Sample
0
.1
.2
.3
.4
.5
0 1 2 3 4 5 6 7 8 9 10 11
Citations
Patents With a Dozen or More Citations
Number
0
200
400
600
800
10 20 30 40 50
Citations

Figure 3: Academic Earmarks



Source: Chronicle of Higher Education

Table 1: Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.
Patent Variables				
Total Patents 90-99	1,907	4,638	7	42,702
Patents per 10,000 Pop., 90-99	2.06	2.11	0.07	17.14
Patents per 10,000 Pop., 85-89	2.76	2.23	0.16	15.65
Citations for Patents granted, 90-99	2,346	6,219	6	57,309
Citations per 10,000 pop, 90-99	2.27	2.65	0.05	25.16
Citations per 10,000 pop., 80-89	1.65	1.70	0	13.57
Scale, Density, Market Structure				
Payroll Employment, 89	392,480	862,483	37,375	9,665,015
UA land area 90 census, sq miles	211	333	15	3,015
MSA Employment / UA Land Area, 90	1,728	689	408	5,021
Scientists & Engineers / UA Land Area, 90	26.15	21.57	1.17	195.04
Establishments per 100,000 Workers, 89	4,425	598	2,667	6,365
Local Universities				
Full Time College Enrollment, 87-89	63,254	116,475	0	1,330,349
College Enrollment / Pop (87-89)	6.8%	5.6%	0%	35.7%
NRC Faculty Rating - S&E, 83	3.75	8.78	0.00	84.60
Local R&D Inputs				
Population with College Degree, 90 (%)	19.5%	6.2%	8.1%	45.4%
Total Acad R&D / 1000 enrollment, 87-89	0.99	1.58	0	11.01
Federal Lab R&D ($1,000) / Fed. Civ. Job, 87-89	1.40	10.81	0	161.39
Private R&D Labs per 1,000 estabs., 87	0.30	0.39	0	2.71
Historical Mix of Industries				
Manufacturing Employment, 89 (%)	14.9%	7.4%	1.8%	46.1%
Services Employment, 89 (%)	25.8%	4.2%	9.8%	44.8%
Construction Employment, 89 (%)	5.4%	1.3%	2.9%	11.0%
Transportation Employment, 89 (%)	4.4%	1.5%	1.6%	11.9%
Wholesale Employment, 89 (%)	4.3%	1.4%	0.7%	9.2%
Retail Employment, 89 (%)	17.7%	1.9%	12.0%	24.8%
FIRE Employment, 89 (%)	6.7%	2.0%	2.7%	16.7%
Local Gov't Employment, 89 (%)	11.7%	4.7%	4.4%	34.5%
Federal Civilian Employment, 89 (%)	2.3%	2.4%	0.3%	20.8%
Historical Mix of Technology				
High-T+A59ech Patents, 80-89 (%)	18.8%	19.5%	0%	88.9%
Chemicals Patents, 80-89 (%)	17.1%	12.8%	0%	76.1%
Computer Patents, 80-89 (%)	5.7%	6.5%	0%	48.2%
Medical Patents, 80-89 (%)	6.4%	6.1%	0%	44.8%
Electrical Patents, 80-89 (%)	14.5%	9.7%	0%	56.3%
Mechanical Patents, 80-89 (%)	24.9%	10.0%	5.6%	62.4%
Other Controls				
Working Age Population, 90 (%)	64.4%	3.1%	53.8%	74.8%
Employment Growth, 80-89 (%)	20.5%	15.5%	-25.8%	77.7%
Far West Region Dummy	0.13	0.33	0	1.00
Great Lakes Region Dummy	0.18	0.38	0	1.00
Mideast Region Dummy	0.12	0.32	0	1.00
New England Region Dummy	0.06	0.23	0	1.00
Plains Region Dummy	0.09	0.29	0	1.00
Southeast Dummy	0.28	0.45	0	1.00
Southwest Dummy	0.11	0.31	0	1.00

Table 1: Descriptive Statistics (cont)

	Mean	Std. Dev.	Min.	Max.
Academic R&D				
pct R&D - all engineering 87-89	9.0%	15.2%	0%	87.1%
pct R&D - all physical sciences 87-89	9.1%	14.4%	0%	81.0%
pct R&D - math & computer science 87-89	4.0%	11.5%	0%	100.0%
pct R&D - all life sciences 87-89	29.5%	31.7%	0%	100.0%
pct R&D - all social sciences 87-89	5.2%	11.3%	0%	100.0%
pct R&D - aerospace engineering 87-89	0.6%	3.0%	0%	35.4%
pct R&D - chemical engineering 87-89	0.8%	2.8%	0%	28.0%
pct R&D - civil engineering 87-89	1.0%	2.5%	0%	21.2%
pct R&D - electrical engineering 87-89	2.1%	5.7%	0%	54.5%
pct R&D - mechanical engineering 87-89	1.3%	3.3%	0%	29.8%
pct R&D - materials science & other 87-89	3.2%	7.5%	0%	57.0%
pct R&D - astronomy 87-89	0.5%	3.1%	0%	34.3%
pct R&D - chemistry 87-89	4.5%	9.0%	0%	81.0%
pct R&D - physics 87-89	3.2%	7.2%	0%	77.5%
pct R&D - other physical sciences 87-89	0.9%	5.1%	0%	73.4%
pct R&D - all geosciences 87-89	5.2%	10.8%	0%	85.9%
pct R&D - mathematics 87-89	0.9%	1.9%	0%	15.6%
pct R&D - computer science 87-89	3.1%	11.0%	0%	100.0%
pct R&D - agricultural science 87-89	4.1%	11.9%	0%	100.0%
pct R&D - biology 87-89	12.9%	17.3%	0%	100.0%
pct R&D - medicine 87-89	11.1%	19.7%	0%	97.6%
pct R&D - other life sciences 87-89	1.4%	5.2%	0%	67.9%
pct R&D - psychology 87-89	1.5%	3.8%	0%	41.3%
pct R&D - economics 87-89	1.1%	3.2%	0%	29.8%
pct R&D - political science & public admin 87-89	0.9%	6.3%	0%	100.0%
pct R&D - sociology 87-89	1.2%	5.1%	0%	62.9%
pct R&D - other social sciences 87-89	1.9%	6.6%	0%	62.8%
pct R&D - interdisciplinary & other science n.e.c.	1.4%	4.0%	0%	36.8%
pct Acad R&D funded by industry 87-89	5.4%	7.7%	0%	56.7%
pct Acad R&D funded by federal govt 87-89	34.4%	31.0%	0%	100.0%
pct Basic Acad R&D funded by fed govt 87-90	24.2%	23.2%	0%	76.7%
pct Applied Acad R&D funded by fed govt 87-89	8.2%	8.6%	0%	50.2%
pct Acad R&D funded by state & local govt 87-89	6.4%	11.0%	0%	62.8%
pct Acad R&D funded by university 87-89	13.2%	16.1%	0%	95.8%
pct Acad R&D funded by other 87-89	5.2%	9.6%	0%	93.6%
Academic Earmarks				
Academic Earmarks, 90	985,302	2,680,645	0	21,000,000
Acad Earmarks, 90 x3 / 1000 enrollment, 87-89	0.043	0.147	0	1.660
DoD Earmarks, 90 x3 / 1000 enrollment, 87-89	0.004	0.044	0	0.714
DoE Earmarks, 90 x3 / 1000 enrollment, 87-89	0.007	0.038	0	0.324
GSA Earmarks, 90 x3 / 1000 enrollment, 87-89	0.003	0.028	0	0.276
NASA Earmarks, 90 x3 / 1000 enrollment, 87-89	0.008	0.103	0	1.660
USDA Earmarks, 90 x3 / 1000 enrollment, 87-89	0.014	0.059	0	0.541

Table 1: Descriptive Statistics (cont)

	Mean	Std. Dev.	Min.	Max.
Instruments				
Jobs / UA Land Area, 1970	1,900	900	600	9,000
UA Land Area, 1970	156	263	12	2,425
1940 MA Population	163,778	581,281	0	8,128,177
Planned Highway Miles, 1947	19	18	0	143
Population with College Degree, 1940 (%)	4.9%	1.8%	1.5%	12.9%

Table 2: Federal Spending on R&D & R&D Earmarks, 1990

	R&D Outlays ($ millions)				Earmarks ($ millions)	
	Total	%	to Universities	%		%
Agriculture Dept.	1,108	1.7	348	3.8	189	46.7
Defense Dept.	37,268	58.6	1,213	13.3	29	7.2
Energy Dept.	5,631	8.9	500	5.5	74	18.2
NASA	6,533	10.3	471	5.2	17	4.2
NIH	7,979	12.6	4,779	52.3	0	0.0
NSF	1,690	2.7	1,321	14.5	0	0.0
Other	3,344	5.3	505	5.5	96	23.6
Total	63,553		9,138		404	

Sources: Chronicle of Higher Education, "Congressional Earmarks for Higher Education," National Science Foundation, "Federal R&D Obligations for Universities & Colleges ," and author's calculations.

Table 3: Main Regressions

	Dependent Variable (in logs):			
	1		2	
			Citation-Weighted	
	Patents / 10,000 Pop.		Patents / 10,000 Pop.	
Independent Variables:	Coef.	Std. Error	Coef.	Std. Error
Scale, Density & Local Market Structure:				
Job Density, 1990†	0.2210	(0.0867) **	0.2168	(0.1102) **
Employment (10,000), 1990†	0.3810	(0.1472) ***	0.5168	(0.1698) ***
Employment Squared†	-0.0494	(0.0177) ***	-0.0563	(0.0207) ***
Establishments per Employee, 1989†	1.5814	(0.3418) ***	1.4209	(0.3913) ***
Local R&D Inputs:				
College Education (percent), 1990†	0.9651	(0.1796) ***	1.0535	(0.2546) ***
College Enrollment / Population, 1990	0.3800	(1.0549)	0.1066	(1.0852)
Academic R&D per student, 1987-89	0.0693	(0.0263) ***	0.0649	(0.0263) **
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0062	(0.0018) ***	0.0078	(0.0022) ***
Private R&D Labs / Establishements, 1989	0.3302	(0.0723) ***	0.4180	(0.0931) ***
Historical Mix of Industries & Technologies:				
Trade Secrets Index (lab weighted)†	0.3685	(0.2440)	0.0847	(0.3463)
Manufacturing Employment (percent), 1989	2.1975	(0.6889) ***	3.2456	(0.8618) ***
Construction Employment (percent), 1989	0.0262	(3.0434)	0.2303	(3.4430)
Transportation Employment (percent), 1989	-1.9562	(2.7307)	-4.1986	(3.2903)
Wholesale Employment (percent), 1989	-3.4096	(3.2619)	-2.6607	(3.8066)
Retail Employment (percent), 1989	-3.7987	(2.2025) *	-1.6757	(2.4747)
Services Employment (percent), 1989	0.1370	(0.9385)	1.3665	(1.1328)
Finance & Real Estate Employment (percent), 1989	0.4545	(1.8083)	0.8305	(1.9221)
Federal Civilian Employment (percent), 1989	-2.1846	(1.4451)	-1.5148	(2.1050)
Local Govt. Employment (percent), 1989	-3.2298	(1.2936) **	-1.7624	(1.3473)
High-Tech Patents (percent), 1980-89	0.8496	(0.1898) ***	0.8629	(0.2075) ***
Chemical Patents (percent), 1980-89	1.7397	(0.3946) ***	1.7287	(0.4850) ***
Computer Patents (percent), 1980-89	3.3144	(0.6041) ***	3.7179	(0.6502) ***
Medical Patents (percent), 1980-89	-0.3111	(0.5845)	0.6480	(0.6293)
Electrical Patents (percent), 1980-89	0.9356	(0.4613) **	1.1309	(0.5932) *
Mechanical Patents (percent), 1980-89	1.1009	(0.4424) **	1.3142	(0.4992) ***
Working Age Population (percent), 1990	2.5352	(1.3948) *	3.5045	(1.6539) **
Employment Growth (percent), 1980-89	0.3579	(0.2427)	0.4414	(0.2741)
Constant	-21.3111	(3.3309) ***	-21.2756	(3.9406) ***
Observations	280		280	
Adjusted R^2	0.7884		0.7968	

Notes: Regressions include 7 dummy variables for census regions. † variable in logs. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 4: Linear and Quadratic Specifications of Scale & Density

Dependent Variable: Patents per 10,000 Population, 1990-99†				
	1	2	3	4
Job Density, 1990†	0.2283	4.3558	0.2210	3.6225
	(0.0890)**	(1.7127)**	(0.0867)**	(1.7879)**
Job Density Squared†		-0.2816		-0.2320
		(0.1161)**		(0.1215)*
Employment (10,000), 1990†	0.0326	0.0350	0.3810	0.3473
	(0.0406)	(0.0398)	(0.1472)***	(0.1530)**
Employment Squared†			-0.0494	-0.0444
			(0.0177)***	(0.0186)**
Optimal Scale (1,000s)	n/a	n/a	473	500
Optimal Density	n/a	2,285	n/a	2,458
n	280	280	280	280
Adjusted R^2	0.7829	0.7860	0.7884	0.7902

Dependent Variable: Citation-Weighted Patents per 10,000 Population, 1990-99†				
	1	2	3	4
Job Density, 1990†	0.2251	6.0982	0.2168	5.2919
	(0.1133)**	(2.4232)**	(0.1102)**	(2.4627)**
Job Density Squared†		-0.4007		-0.3462
		(0.1639)**		(0.1668)**
Employment (10,000), 1990†	0.1197	0.1231	0.5168	0.4666
	(0.0474)**	(0.0460)***	(0.1698)***	(0.1729)***
Employment Squared†			-0.0563	-0.0488
			(0.0207)***	(0.0211)**
Optimal Scale (1,000s)	n/a	n/a	985	1,192
Optimal Density	n/a	2,017	n/a	2,086
n	280	280	280	280
Adjusted R^2	0.7968	0.7968	0.7968	0.8002

Notes: Regressions include all the other variables included in Table 3. †: log of the variable. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 5: Elasticities at Mean of Independent variables

	Dependent Variable (in logs):					
	Patents / 10,000 Pop.			Citation-Weighted Patents / 10,000 Pop.		
Independent Variables:	Coef.	Std. Error		Coef.	Std. Error	
Local R&D Inputs:						
College Education (percent), 1990	0.9651	(0.1796)	***	1.0535	(0.2546)	***
Academic R&D per student, 1987-89	0.0686	(0.0260)	***	0.0642	(0.0260)	**
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0086	(0.0025)	***	0.0109	(0.0031)	***
Private R&D Labs / Establishements, 1989	0.1003	(0.0220)	***	0.1269	(0.0283)	***
Historical Mix of Industries & Technologies:						
Manufacturing Employment (percent), 1989	0.3279	(0.1028)	***	0.4842	(0.1286)	***
Local Govt. Employment (percent), 1989	-0.3792	(0.1519)	**	-0.2069	(0.1582)	
High-Tech Patents (percent), 1980-89	0.1596	(0.0356)	***	0.1620	(0.0390)	***
Chemical Patents (percent), 1980-89	0.2982	(0.0676)	***	0.2963	(0.0831)	***
Computer Patents (percent), 1980-89	0.1887	(0.0344)	***	0.2116	(0.0370)	***
Medical Patents (percent), 1980-89	-0.0200	(0.0375)		0.0416	(0.0404)	
Electrical Patents (percent), 1980-89	0.1353	(0.0667)	**	0.1635	(0.0858)	*
Mechanical Patents (percent), 1980-89	0.2741	(0.1101)	**	0.3272	(0.1243)	***

Notes: Based on regressions reported in Table 3. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 6: Elasticities at Mean - Fields of Science & Other (select) Variables

	Dependent Variable (in logs):					
	Patents / 10,000 Pop.			Citation-Weighted Patents / 10,000 Pop.		
Independent Variables:	Coef.	Std. Error		Coef.	Std. Error	
Local R&D Inputs:						
Academic R&D per student, 1987-89	0.0748	(0.0287)	***	0.0708	(0.0282)	**
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0078	(0.0026)	***	0.0101	(0.0032)	***
Private R&D Labs / Establishements, 1989	0.0940	(0.0215)	***	0.1171	(0.0265)	***
Coarse Fields of Science:						
Pct Acad R&D - All Engineering Sciences, 87-89	-0.0144	(0.0152)		-0.0137	(0.0181)	
Pct Acad R&D - All Physical Sciences, 87-89	0.0085	(0.0153)		0.0199	(0.0207)	
Pct Acad R&D - All Earth Sciences, 87-89	-0.0055	(0.0087)		0.0069	(0.0108)	
Pct Acad R&D - All Mathematical Sciences, 87-89	-0.0206	(0.0066)	***	-0.0273	(0.0101)	***
Pct Acad R&D - All Life Sciences, 87-89	-0.0593	(0.0305)	*	-0.0622	(0.0309)	**
Pct Acad R&D - All Social Sciences, 87-89	-0.0091	(0.0113)		-0.0060	(0.0125)	
Observations	280			280		
Adjusted R^2	0.7923			0.8040		

Notes: The excluded field is interdisciplinary & other science. Regressions also include all the other variables reported in Table 3. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 7: Elasticities at Mean - Fields of Science & Other (select) Variables

	Dependent Variable (in logs):					
	Patents / 10,000 Pop.			Citation-Weighted Patents / 10,000 Pop.		
Independent Variables:	Coef.	Std. Error		Coef.	Std. Error	
Local R&D Inputs:						
Academic R&D per student, 1987-89	0.0815	(0.0331)	**	0.0802	(0.0328)	**
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0080	(0.0028)	***	0.0100	(0.0034)	***
Private R&D Labs / Establishements, 1989	0.0914	(0.0223)	***	0.1151	(0.0287)	***
Narrow Fields of Science:						
Pct Acad R&D - Aeronautical Engineering, 87-89	-0.0044	(0.0044)		-0.0045	(0.0046)	
Pct Acad R&D - Chemical Engineering, 87-89	0.0020	(0.0070)		0.0073	(0.0073)	
Pct Acad R&D - Civil Engineering, 87-89	-0.0036	(0.0110)		0.0004	(0.0123)	
Pct Acad R&D - Electrical Engineering, 87-89	-0.0060	(0.0127)		0.0016	(0.0129)	
Pct Acad R&D - Mechanical Engineering, 87-89	0.0021	(0.0120)		0.0003	(0.0134)	
Pct Acad R&D - Materials Science & Other, 87-89	-0.0010	(0.0100)		-0.0123	(0.0132)	
Pct Acad R&D - Astronomy, 87-89	-0.0014	(0.0047)		-0.0023	(0.0059)	
Pct Acad R&D - Chemistry, 87-89	0.0117	(0.0123)		0.0135	(0.0154)	
Pct Acad R&D - Physics, 87-89	-0.0012	(0.0090)		0.0039	(0.0114)	
Pct Acad R&D - Other Physical Sciences, 87-89	0.0010	(0.0042)		0.0020	(0.0063)	
Pct Acad R&D - Earth Sciences, 87-89	-0.0029	(0.0091)		0.0091	(0.0115)	
Pct Acad R&D - Mathematics, 87-89	-0.0066	(0.0105)		-0.0090	(0.0161)	
Pct Acad R&D - Computer Science, 87-89	-0.0149	(0.0053)	***	-0.0195	(0.0080)	**
Pct Acad R&D - Agricultural Sciences, 87-89	-0.0151	(0.0116)		-0.0182	(0.0103)	*
Pct Acad R&D - Biology, 87-89	-0.0184	(0.0182)		-0.0214	(0.0173)	
Pct Acad R&D - Medicine, 87-89	-0.0293	(0.0180)		-0.0315	(0.0192)	
Pct Acad R&D - Other Life Sciences, 87-89	-0.0007	(0.0065)		-0.0011	(0.0084)	
Pct Acad R&D - Psychology, 87-89	-0.0171	(0.0083)	**	-0.0167	(0.0099)	*
Pct Acad R&D - Economics, 87-89	-0.0007	(0.0125)		0.0034	(0.0147)	
Pct Acad R&D - Political Science & Public Adm., 87-89	-0.0020	(0.0014)		-0.0030	(0.0017)	*
Pct Acad R&D - Sociology, 87-89	0.0067	(0.0036)	*	0.0075	(0.0059)	
Pct Acad R&D - Other Social Sciences, 87-89	-0.0115	(0.0087)		-0.0087	(0.0087)	
Observations	280			280		
Adjusted R^2	0.7846			0.7968		

Notes: The excluded field is interdisciplinary & other science. Regressions also include all the other variables reported in Table 3. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 8: Elasticities at Mean - Sources of Academic R&D Funding

	Dependent Variable (in logs):							
	1		2		3		4	
	Patents / 10,000 Pop.		Patents / 10,000 Pop.		Citation-Weighted Patents / 10,000 Pop.		Citation-Weighted Patents / 10,000 Pop.	
Independent Variables:	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
Local R&D Inputs:								
Academic R&D per student, 1987-89	0.0795	(0.0272) ***	0.0873	(0.0274) ***	0.0786	(0.0274) ***	0.0845	(0.0273) ***
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0079	(0.0024) ***	0.0081	(0.0024) ***	0.0100	(0.0030) ***	0.0103	(0.0030) ***
Private R&D Labs / Establishements, 1989	0.1061	(0.0203) ***	0.1024	(0.0198) ***	0.1311	(0.0262) ***	0.1289	(0.0257) ***
Source of R&D Funding:								
Pct Acad R&D - Federal Funds, 87-89	-0.1353	(0.0359) ***			-0.1390	(0.0390) ***		
Pct Acad R&D - State & Local Funds, 87-89	-0.0219	(0.0142)	-0.0042	(0.0194)	-0.0205	(0.0191)	0.0025	(0.0240)
Pct Acad R&D - Industry Funds, 87-89	-0.0074	(0.0195)	0.0378	(0.0108) ***	-0.0005	(0.0240)	0.0484	(0.0143) ***
Pct Acad R&D - Other Sources, 87-89	0.0383	(0.0109) ***	-0.0633	(0.0462)	0.0488	(0.0145) ***	-0.0619	(0.0498)
Pct Acad R&D - Federal Funds - Basic R&D			-0.0633	(0.0462)			-0.0619	(0.0498)
Pct Acad R&D - Federal Funds - Applied R&D			-0.0772	(0.0427) *			-0.0789	(0.0442) *
Observations	280		280		280		280	
Adjusted R^2	0.8035		0.8039		0.8108		0.8108	

Notes: The excluded share is internal funding. Regressions include all the other variables reported in Table 3. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 9: Elasticities at Mean - Academic Earmarks

	Dependent Variable (in logs):											
	1			2			3			4		
							Citation-Weighted			Citation-Weighted		
	Patents / 10,000 Pop.			Patents / 10,000 Pop.			Patents / 10,000 Pop.			Patents / 10,000 Pop.		
Independent Variables:	Coef.	Std. Error		Coef.	Std. Error		Coef.	Std. Error		Coef.	Std. Error	
Local R&D Inputs:												
Academic R&D per student, 1987-89	0.0832	(0.0279)	***	0.0742	(0.0271)	***	0.0826	(0.0284)	***	0.0739	(0.0270)	***
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0076	(0.0024)	***	0.0077	(0.0024)	***	0.0097	(0.0030)	***	0.0098	(0.0030)	***
Private R&D Labs / Establishements, 1989	0.1033	(0.0203)	***	0.1026	(0.0207)	***	0.1281	(0.0263)	***	0.1293	(0.0265)	***
Source of R&D Funding:												
Pct Acad R&D - Federal Funds, 87-89	-0.1385	(0.0360)	***	-0.1429	(0.0368)	***	-0.1415	(0.0398)	***	-0.1522	(0.0392)	***
Pct Acad R&D - State & Local Funds, 87-89	-0.0219	(0.0142)		-0.0222	(0.0146)		-0.0208	(0.0190)		-0.0183	(0.0196)	
Pct Acad R&D - Industry Funds, 87-89	-0.0071	(0.0193)		-0.0059	(0.0191)		-0.0003	(0.0238)		0.0006	(0.0241)	
Pct Acad R&D - Other Sources, 87-89	0.0373	(0.0111)	***	0.0372	(0.0112)	***	0.0471	(0.0147)	***	0.0474	(0.0148)	***
Federal Academic Earmarks per student#	-0.0068	(0.0057)		-0.0464	(0.0131)	***	-0.0154	(0.0070)	**	-0.0544	(0.0214)	**
Dod Academic Earmarks per student#				0.0050	(0.0014)	***				0.0070	(0.0021)	***
DoE Academic Earmarks per student#				0.0050	(0.0033)					0.0063	(0.0047)	
GSA Academic Earmarks per student#				0.0010	(0.0017)					0.0021	(0.0024)	
NASA Academic Earmarks per student#				0.0080	(0.0026)	***				0.0078	(0.0040)	*
USDA Academic Earmarks per student#				0.0186	(0.0079)	**				0.0109	(0.0094)	
Observations	280			280			280			280		
Adjusted R^2	0.8033			0.8054			0.8123			0.8120		

Notes: The excluded share is internal funding. #: Earmarks in 1990 x 3, divided by full-time enrollment in the MA in 1987-89. Regressions include all the other variables reported in Table 3. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 10: Regressions Using Density of Scientists & Engineers

	Dependent Variable (in logs):			
	Patents / 10,000 Pop.		Citation-Weighted Patents / 10,000 Pop.	
Independent Variables:	Coef.	Std. Error	Coef.	Std. Error
Scale, Density & Local Market Structure:				
S&E Job Density, 1990†	0.2025	(0.0705) ***	0.2646	(0.0829) ***
Employment (10,000), 1990†	0.3204	(0.1473) **	0.4337	(0.1684) **
Employment Squared†	-0.0455	(0.0174) ***	-0.0509	(0.0202) **
Establishments per Employee, 1989†	1.4616	(0.3312) ***	1.2923	(0.3740) ***
Local R&D Inputs:				
College Education (percent), 1990†	0.8928	(0.1852) ***	0.9483	(0.2611) ***
College Enrollment / Population, 1990	0.4926	(1.0969)	0.2147	(1.0739)
Academic R&D per student, 1987-89	0.0681	(0.0259) ***	0.0619	(0.0254) **
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0043	(0.0018) **	0.0056	(0.0021) ***
Private R&D Labs / Establishements, 1989	0.2730	(0.0714) ***	0.3472	(0.0919) ***
Observations	278		278	
Adjusted R^2	0.7973		0.8097	

Notes: Except for the density variable, the regression is identical to the one reported in Table 3. † variable in logs. The dependent variable includes all patents/patent citations over the years 1990-99. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 11: Main Regressions, with Lagged Dependent Variable

	Dependent Variable (in logs):			
	Patents / 10,000 Pop.		Citation-Weighted Patents / 10,000 Pop.	
Independent Variables:	Coef.	Std. Error	Coef.	Std. Error
Scale, Density & Local Market Structure:				
Job Density, 1990†	0.1370	(0.0653) **	0.1856	(0.1005) *
Employment (10,000), 1990†	0.1430	(0.0950)	0.4113	(0.1237) ***
Employment Squared†	-0.0253	(0.0114) **	-0.0469	(0.0159) ***
Establishments per Employee, 1989†	0.1075	(0.2272)	0.5408	(0.3145) *
Local R&D Inputs:				
College Education (percent), 1990†	0.4692	(0.1173) ***	0.7707	(0.1510) ***
College Enrollment / Population, 1990	-0.1786	(0.7388)	-0.6548	(0.8651)
Academic R&D per student, 1987-89	0.0432	(0.0161) ***	0.0400	(0.0184) **
Federal Lab R&D / Fed Civ Jobs, 1987-89	0.0035	(0.0013) ***	0.0058	(0.0019) ***
Private R&D Labs / Establishements, 1989	0.0378	(0.0558)	0.2318	(0.0735) ***
Other				
Patent intensity, 1980s† #	0.6877	(0.0527) ***	0.3869	(0.0667) ***
Constant	-7.5713	(2.2655) ***	-8.3142	(3.4159) **
Observations	280		254	
Adjusted R^2	0.8889		0.8582	

Notes: Regressions are identical to those reported in Table 3, except for the addition of the lagged dependent variable. The dependent variable includes all patents/patent citations over the years 1990-99. † variable in logs. # The average of patents per 10,000 population over 1985-89 or citation-weighted patents per 10,000 population over 1980-89. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent.

Table 12a: Instrumenting for Job Density (2SLS)

	Dependent Variable (in logs):					
	Patents / 10,000 Pop.			Citation-Weighted Patents / 10,000 Pop.		
	1	2	3	4	5	6
Second Stage:						
Job Density, 1990[†] [#]	0.437	0.2900	0.2931	0.5477	0.3800	0.3603
	(0.2290)*	(0.1379)***	(0.1091)***	(0.2861)*	(0.1807)**	(0.1426)**
Other Exogenous Variables	No	Yes	Yes	No	Yes	Yes
Other (Potentially) Endogenous Variables	No	No	Yes	No	No	Yes
Over-identification Test (*p* values)	0.9651	0.8642	0.5003	0.5718	0.4636	0.6950
Endogeneity Test (*p* values)	0.2254	0.8470	0.3531	0.1282	0.5987	0.2456
First Stage:						
F Statistic	51.06	61.53	62.75	51.06	61.53	62.75
Partial R^2	0.4707	0.5347	0.5316	0.4707	0.5347	0.5316
Observations	224	224	224	224	224	224

Notes: [†] Variable in logs. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent

Variable is instrumented with the square of MA job density (in logs) in 1970 and a dummy variable for the significant presence of hills and mountains

Table 12b: Instrumenting for Employment (2SLS)

	Dependent Variable (in logs):					
					Citation-Weighted	
		Patents / 10,000 Pop.			Patents / 10,000 Pop.	
	1	2	3	4	5	6
Second Stage:						
Employment (10,000), 1990†#	0.4382	0.4465	0.3057	0.4284	0.4798	0.3302
	(0.3377)	(0.2166)**	(0.1970)	(0.4041)	(0.2372)**	(0.2206)
Employment Squared†#	-0.0302	-0.0570	-0.0393	-0.0931	-0.0527	-0.0339
	(0.0445)	(0.0262)**	(0.0239)*	(0.0689)	(0.0286)*	(0.0267)
Other Exogenous Variables	No	Yes	Yes	No	Yes	Yes
Other (Potentially) Endogenous Variables	No	No	Yes	No	No	Yes
Over-identification Test (*p* values)	0.4840	0.2309	0.4750	0.4859	0.1020	0.2600
Endogeneity Test (*p* values)	0.0008	0.9757	0.8778	0.0000	0.5902	0.4740
First Stage:						
F Statistic	222.78	192.87	185.75	222.78	192.87	185.75
	286.87	214.93	212.72	286.87	214.93	212.72
Partial R^2	0.7695	0.7567	0.7567	0.7695	0.7567	0.7567
	0.8209	0.8093	0.8096	0.8209	0.8093	0.8096
Observations	277	277	277	277	277	277

Notes: † Variable in logs. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. *** significant at 1 percent

Variables are instrumented with the log of MA population in 1940 and the log of planned highway+B75 miles in 1947, and the square of those variables.

Table 12c: Instrumenting for Human Capital (2SLS)

	Dependent Variable (in logs):					
	Patents / 10,000 Pop.			Citation-Weighted Patents / 10,000 Pop.		
	1	2	3	4	5	6
Second Stage:						
College Education (percent), 1990[†#]	1.9758	2.3566	2.0177	2.5871	2.7188	2.1731
	(0.2349)***	(0.7458)***	(0.7404)***	(0.2665)***	(0.8677)***	(0.8368)***
Other Exogenous Variables	No	Yes	Yes	No	Yes	Yes
Other (Potentially) Endogenous Variables	No	No	Yes	No	No	Yes
Over-identification Test (*p* values)[a]	N/A	N/A	N/A	N/A	N/A	N/A
Endogeneity Test (*p* values)	0.0049	0.0237	0.0777	0.0011	0.0142	0.0821
First Stage:						
F Statistic	186.95	13.17	12.04	186.95	13.17	12.04
Partial R^2	0.4775	0.0728	0.0653	0.4775	0.0728	0.0653
Observations	279	279	279	279	279	279

Notes: † Variable in logs. Standard errors are corrected for heteroskedasticity. * significant at 10 percent. ** significant at 5 percent. ***
Variable is instrumented with the log of the share of the MA population with a college degree in 1940.
a Over-identification tests, using the annual inches of rainfall (in logs) as a second instrument, do not reject the null hypothesis of instrument endogeneity. The F statistics in the first-stage regressions in columns 3 and 6, however, are just below 10. The second stage results are similar to those reported here.

Table 13: Tests for Spatial Dependence (*P* values reported)

Test (Null Hypothesis)	Dependent Variable (in logs): Spatial Error: Patents per 10,000 Pop.	Spatial Error: Citation-Wtd. Patents per 10,000 Pop.	Spatial Lag: Patents per 10,000 Pop.	Spatial Lag: Citation-Wtd. Patents per 10,000 Pop.
Moran's I (λ = 0)	0.1583	0.2676	—	—
Lagrange Multiplier (λ = 0)	0.2120	0.2698	—	—
Robust Lagrange Multiplier (λ = 0)	0.6872	0.8940	—	—
Lagrange Multiplier (ρ = 0)	—	—	0.1889	0.2158
Robust Lagrange Multiplier (ρ = 0)	—	—	0.5654	0.5644

Notes: N = 280. Regressions use the specifications reported in Table 3. The dependent variable includes all patents / patent citations over the years 1990-99. Moran's I is based on standardized z values that follow a normal distribution. The Lagrange multiplier tests are distributed $\chi 2$ with critical levels 3.84 ($p = 0.05$).