Is Academic Science Driving a Surge in Industrial Innovation? Evidence from Patent Citations

Lee Branstetter

Discussion Paper No. 28

Lee Branstetter

Associate Professor
Columbia Business School

Discussion Paper Series
APEC Study Center
Columbia Business School

May 2004

IS ACADEMIC SCIENCE DRIVING A SURGE IN INDUSTRIAL INNOVATION? EVIDENCE FROM PATENT CITATIONS*

Lee Branstetter
Associate Professor
Columbia Business School
815 Uris Hall
3022 Broadway
New York, NY 10027
and NBER

lgb2001@columbia.edu
This Version: January 2003

ABSTRACT

What is driving the remarkable increase over the last decade in the propensity of patents to cite academic science? Does this trend indicate that stronger knowledge spillovers from academia have helped power the surge in innovative activity in the U.S. in the 1990s? This paper seeks to shed light on these questions by using a common empirical framework to assess the relative importance of various alternative hypotheses in explaining the growth in patent citations to science. My analysis supports the notion that the nature of U.S. inventive activity has changed over the sample period, with an increased emphasis on the use of the knowledge generated by university-based scientists in later years. However, the concentration of patent-to-paper citation activity within what I call the "bio nexus" suggests that much of the contribution of knowledge spillovers from academia may be largely confined to bioscience-related inventions.

* **Acknowledgements**: I thank seminar participants at the NBER Summer Institute, Columbia Business School, the UCLA Anderson Graduate School of Management, Hitotsubashi University, the University of Tokyo, and the NBER-CREST-CRIW Zvi Griliches Memorial Conference for useful comments and suggestions. I am grateful to Pierre Azoulay, Rebecca Henderson, Adam Jaffe, and Joshua Lerner for detailed comments on an earlier draft. I also wish to thank a number of academic scientists and industrial R&D managers for providing me with their insights into the process by which knowledge flows from academia to industry. By prior agreement, they will remain anonymous. I am indebted to Masami Imai, Hiau-Looi Kee, Changxiu Li, Kaoru Nabeshima, and, especially, Yoshiaki Ogura for excellent research assistance. I would like to thank Tony Breitzman and Francis Narin of CHI-Research, Adam Jaffe, and Marie and Jerry Thursby for their help in obtaining the data used in this study. The Institute of Economic Research at Hitotsubashi University provided a hospitable environment in which part of this paper was written – I thank the faculty and staff for their support. This project was funded by grants from the University of California Industry-University Cooperative Research Program, the NBER Project on Industrial Technology and Productivity, the Japan Foundation Center for Global Partnership, and the National Science Foundation. All views expressed in this paper, however, are the author's own opinions.

I. Introduction

Recent research points to an evident surge in innovative activity in the United States over the past fifteen years.[1] This is suggested by, among other things, a sharp rise in patent applications and patent grants that started in the late 1980s and has persisted through the end of the 1990s – a rise that has outpaced, by a considerable margin, increases in public and private R&D spending. While a large fraction of U.S. patent grants are awarded to foreign inventors, the fraction obtained by domestic inventors has risen – and this fraction has risen particularly rapidly in fields where patenting has grown most sharply. The recent patent surge could potentially be explained by an increase in the propensity of Americans to patent inventions, rather than an increase in the productivity of American research and development, but the recent research of Kortum and Lerner [1998, 2000, 2003] strongly suggests that recent trends in patenting and related data are more consistent with the latter interpretation. If this conclusion is correct, then it could help explain the widely observed increase in U.S. TFP growth in recent years.[2]

But if American R&D productivity has increased, then that raises the question of what factors are driving the increase.[3] This paper attempts to assess the importance of one possible contributing factor – increased knowledge spillovers from U.S.-based academic science. In essence, this paper is an attempt to explain the phenomenon graphed out in Figure I. This figure shows that citations made by patents granted in the United States to articles in the scientific literature increased very rapidly from the mid 1980s through the late 1990s.[4] Over this period, the number of patents granted by the U.S. Patent and Trademark Office to U.S. residents more than doubled, real R&D expenditures in the United States rose by almost 40%, and global output of scientific articles increased by about 13%, but patent citations to science *increased more than 13 times*.[5] Many at the National Science Foundation and other U.S. science policy agencies find this graph extremely interesting, because it seems to suggest – at least in some broad sense – that

---

[1] See Jaffe and Lerner [forthcoming], Kortum and Lerner [1998], Kortum and Lerner [2000], and Kortum and Lerner [2003].

[2] See Gordon [2000] and DeLong [2001].

[3] The work of Kortum and Lerner [2000] has stressed the potential role of venture capital-linked firms in improving U.S. R&D output.

[4] This graph does not break down growth in citations by the nationality of the inventor, but data from the 2002 *National Science and Engineering Indicators* shows that the majority of these citations are made by domestic patent applicants, and U.S.-based academic science is disproportionately likely to be cited. The fraction of citations to science made to U.S. authors has increased over this period. See also Narin et. al. [1997] and Hicks et. al. [2001].

[5] These data come from the 2002 edition of the *National Science and Engineering Indicators*. The data on scientific article output may understate the growth in articles, but even a substantial correction of the official statistics would leave the basic message of Figure 1 essentially unchanged.

academic science and industrial technology are "closer" than they used to be. This could mean that publicly funded science is generating more spillovers to industrial innovation than in the past.[6] This, in turn, may have contributed in important ways to the apparent surge of innovative activity in the United States in the 1990s.

This positive interpretation of recent trends in the data is influenced by the theoretical contributions of Evenson and Kislev [1976] and the more recent analysis their work inspired, such as Adams [1990] and Kortum [1997]. In this general class of models, applied research is a search process that eventually exhausts the technological opportunities within a particular field. However, basic science can open up new "search distributions" for applied researchers, raising the productivity and the level of applied research effort – at least temporarily. Viewed through this theoretical lens, the concurrence of rapid growth in U.S. private R&D expenditures, even more rapid growth in patenting, mounting evidence of an acceleration in TFP growth, and still more rapid growth in the intensity with which U.S. patents cite academic science would all seem to suggest a response to new technological opportunities created by academic research. Not surprisingly, other advanced industrial nations are deliberately trying to foster closer connections between university-based scientific research and industrial R&D in conscious imitation of the "U.S. model."

However, increasingly strong knowledge spillovers from academic science to industrial R&D are only one of several factors that could be driving the changes illustrated in Figure I. Furthermore, even if such knowledge spillovers are growing in strength, this could be happening in a number of different ways, which have different implications for public policy. A little thought and a cursory reading of the recent literature generate at least four alternative hypotheses that could explain the recent trends in the data. The first is the ***"increasing scientific fertility"*** hypothesis, which posits that more recent cohorts of scientific papers contain more discoveries that are directly applicable to industrial research and development, and that this trend holds across many fields of science. Under this hypothesis, knowledge spillovers from academia to industry are increasing primarily because of a qualitative change in the nature of the science being conducted at universities.[7]

---

[6] This interpretation has been stressed in recent editions of the *National Science and Engineering Indicators* and in the recent work of Narin et. al. [1997].

[7] I will note that here and elsewhere, I am being a bit loose in my use of the term "knowledge spillover." The knowledge flows from academia to industry are only pure spillovers to the extent that industrial inventors receive them for free. In fact, conversations with industry-based R&D managers suggest that investments on the part of the firm (of various kinds) are necessary in order to effectively learn from these knowledge flows – so that they are not pure spillovers. See Cohen and Levinthal [1988], Zucker et. al. [1998], and Cockburn, Henderson, and Stern [1999].

The second is the ***"changing methods of invention"*** hypothesis, which posits that industrial inventors have changed the way they create new technology. The new approach to R&D draws more heavily on academic science than in the past, though it does not necessarily draw exclusively on the most recently published articles. This would be reflected in an increasing propensity for more recent cohorts of patents across a wide range of technical fields to cite science. Now, this increased propensity for more recent patents to cite science could very well reflect a response by firms to new "technological opportunities" generated by academic scientific breakthroughs. The point being stressed is that it is the inventors themselves who are generating the increased citations as they alter the direction and nature of their R&D programs to probe the new opportunities for industrial research created by basic science. Like the first hypothesis, this implies that knowledge spillovers from academic science are increasing over time, but the mechanism driving this increase is different.

The third is the ***"changing composition of invention"*** hypothesis, which posits that invention in certain areas of technology has been closely linked to science for some time, and, likewise, some fields of science have always been frequently cited by industrial patents. Under this hypothesis, there has been disproportionate growth in patenting in frequently citing patent classes. Similarly, growth in academic publications has been biased towards those fields of science which have historically been more closely linked to industrial R&D. In other words, at the level of individual technology classes and scientific fields, there has been little change in the relationship between science and technology *per se* – rather there has been a change in the distribution of patents and papers that generates the observed increase in citations. A variant of this hypothesis notes that there has been rapid growth in patenting by universities, and that this change in the *composition of inventors* might also contribute to the growth in patent citations to science.

Strongly biased growth in frequently citing patent classes and frequently cited fields of science could itself reflect a response by both industrial inventors and academic scientists to the "technological opportunities" created by a series of fundamental scientific breakthroughs. In fact, one might find within this "nexus" of patent classes and scientific fields evidence of changing methods of invention and/or increases in scientific fertility, such that the intensity of interaction between science and invention actually grows over time. The point being stressed in this "changing composition" hypothesis is that the new technological opportunities, if they exist, are quite specific to a small number of technical and scientific fields, and one does not observe a broad-based change across fields of technology or fields of science that is consistent with substantially changing methods of invention or substantially increased scientific fertility.

The fourth hypothesis is the ***"attorney-driven"*** hypothesis, which posits that the change in patent citations is entirely driven by changes in citations practices. For various strategic reasons connected to the desire to impress patent examiners, the fear of subsequent litigation, or both, patent lawyers have instructed their clients to increase the number of citations made to the scientific literature. The increasing availability of data on the scientific prior art in electronic form has lowered the costs of such citations, further contributing to their growth. This hypothesis, in its extreme version, suggests that little can be learned about the changing relationship between science and technology from patent citation data.

These hypotheses are not mutually exclusive, but they have quite different implications for the appropriate interpretation of the growth in patent citations to papers. In order to understand what Figure I really means, how it relates (or not) to the recent American innovation surge, and what the appropriate policy response is, it is necessary to sort out the relative importance of these hypotheses in explaining the trend illustrated in that graph.

The rest of the paper is largely devoted to an examination of the relative importance of these hypotheses within a common empirical framework. I find that aggregate trends in the data are largely explained by a combination of the "composition hypothesis" and the "changing methods of invention" hypothesis. To a surprising extent, the measured increase in patent citations to papers is localized within a relatively narrow set of technologies and scientific fields related to biotechnology that I will term the "bio nexus." Patenting and publication in these fields has grown over time, and inventors working in these technologies have substantially increased the extent to which they build on science. Citations to science have also increased outside the bio nexus, and the relative change over time has been substantial – but the total numbers of citations outside the bio nexus remain relatively small. In the raw data, there is also ample evidence of a dramatic "attorney-driven" increase in academic citation in the mid-1990s. However, controlling for this legally-driven increase does not qualitatively affect the relative importance of changing composition and changing methods of invention. Key aspects of these conclusions are consistent with other recent papers in this area.

The next section places my approach in the context of the emerging literature on the interaction between academic science and industrial invention. I go on to describe the empirical framework employed in this paper, and report my main findings. In the concluding section, I outline some policy implications of my results and directions for future research. The main message of this paper is that increased knowledge flows from academia may have contributed significantly to the innovation surge reflected in the U.S. patent statistics, but most of that impact is confined to a narrow locus of technologies and scientific fields.

II. The Link Between Academic Science and Industrial Innovation

Historical Perspective

From their inception, publicly supported universities in the U.S. were focused on training students in the "practical arts."[8] In the late 19th and 20th centuries, the search for commercial applications of the preceding decades' scientific discoveries led to the early creation within American universities of new engineering disciplines, including chemical engineering, electrical engineering, and aeronautical engineering. However, progress at the scientific frontier was still dominated by European institutions until the cataclysm of World War II.

The large U.S. postwar investment in basic research, much of it concentrated in universities, and the mass migration of leading European scientists to the United States quickly established America as the leading center of frontier scientific research [Rosenberg and Nelson, 1994]. The infusion of federal funds was predicated on the notion that investment in basic science would eventually lead to useful technological invention for use in both industry and in national defense. However, early attempts to assess the strength of this connection in the postwar era suggested that relationship between "frontier" academic science and industrial invention, while obviously important, was neither close nor direct.[9]

Lessons from the Recent Literature

Drawing upon a wide range of data sources and methodological approaches, the recent economics literature suggests that the linkage between frontier science and industrial technology is stronger and more direct than in the past.[10] Case studies, manager interviews, and surveys have been used to assess the magnitude of this impact, the channels through which it flows, and changes in these factors over time.[11] These studies suggest that firms perceive academic research to be an important input into their own research process, though this importance differs widely

---

8 Rosenberg and Nelson [1994] provide an excellent study of the history of interaction between American universities and industry.

9 See, for example, Derek De Solla Price [1965] and Lieberman [1978]. This view was generally supported by the Defense Department's ambitious "Project Hindsight" study of the impact of basic scientific research on weapons development, which concluded that the primary impact came not from science at the research frontier, but instead from "packed-down, thoroughly understood, carefully taught old science," such as that typically presented in textbooks or university courses. See Sherwin and Isenson [1967], from which the quoted phrase is taken, for a review of Project Hindsight.

10 For a comprehensive literature review that covers relevant research beyond the economics journals, see Agrawal [2001].

11 Important recent studies relying primarily on case study techniques and surveys include Mansfield [1995], Cohen et. al. [1994], Faulkner and Senker [1995], Gambardella [1995], and Agrawal and Henderson [2002].

across firms and industries.[12] A second stream of recent research has undertaken quantitative studies of knowledge spillovers from academic research. Jaffe [1989] and Adams [1990] were early contributors to this literature. More recently, Jaffe et. al. [1993, 1996, 1998] have used data on university patents and citations to these patents to quantify knowledge spillovers from academic science.[13] While patenting by universities has increased substantially in the United States over the last twenty years, there is evidence that as the number of university patents has grown, the marginal quality of those patents has declined.[14]

A related stream of research has undertaken quantitative analysis of university-industry research collaboration. Contributors include Zucker et. al. [1998] and Cockburn and Henderson [1998, 2000]. A number of papers in this literature have studied "start-up" activity related to academic science or academic scientists, such as Zucker et. al. [1998] and Audretsch and Stephan [1996]. Finally, several recent studies have examined university licensing of university generated inventions, such as Barnes et al. [1998], Mowery et. al. [1998], Thursby and Thursby [2002], Shane [2000, 2001], and Lach and Schankerman [2003]. While the counts of licensed inventions have grown over time, there is also evidence that, like patents, the marginal value of licenses has declined as their number has increased [Thursby and Thursby, 2002]. Furthermore, this stream of literature suggests that inventions generated by universities are typically quite "embryonic" – bringing such inventions to the market requires extensive additional investment by private firms.

Using Patent Citations to Academic Science as Measures of Knowledge Spillovers

This paper will use patent citations to academic papers to measure knowledge spillovers between academic science and industrial R&D.[15] As indicators of knowledge spillovers from academia to the private sector, these data have a number of advantages. The academic promotion system creates strong incentives for academic scientists, regardless of discipline, to publish all research results of scientific merit. As a consequence, the top-ranked research universities generate thousands of academic papers each year. Similarly, inventors have an incentive to patent their useful inventions, and a legal obligation under U.S. patent law to make appropriate citations to the prior art – including academic science.

[12] While the channels by which firms absorb the results of academic research vary across industries, the Cohen et. al. [1994] study suggests that the formal scientific literature is, on average, an important channel.

[13] Barnes, Mowery, and Ziedonis [1998] and Mowery, Nelson, Sampat, and Ziedonis [1998] have undertaken a similar study for a smaller number of universities.

[14] See Jaffe, Trajtenberg, and Henderson [1998] and Hicks et al. [2001].

[15] In doing so, I am building on the work of Francis Narin and his collaborators, who have pioneered the use of these data in large-sample "bibliometric" analysis. See Narin et al. [1997] and Hicks et al. [2001] for recent examples of this work.

The recent research discussed in previous paragraphs indicates that, in response to the Bayh-Dole Act and other public policy measures, universities have increased the extent to which they patent the research of university-affiliated scientists. They have also increased the extent to which they license these patented technologies to private firms. Nevertheless, it is clear to observers that only a *tiny fraction* of the typical research university's commercially relevant research output is ever patented, and only a fraction of this set of patents is ever licensed.[16] To illustrate this, Figure II shows the trends over the 1988-1997 period in several alternative indices of university research output and knowledge spillovers for one of the university systems in my data set, the University of California, which includes nine separately managed campuses and a number of affiliated laboratories. The figure graphs university patents by issue year (patents), invention disclosures by year of disclosure filing (invention disclosures), new licenses of university technology by date of contract (licenses), the number of citations to previous university patents by issue year of the citing patent (citations to UC patents), and the number of citations to UC-generated academic papers by issue year of the citing patent (citations to UC papers). Clearly, citations to papers are far more numerous than any other indicator. This figure suggests that patent citations to academic papers may provide a much broader window through which to observe knowledge spillovers from academic science to inventive activity than any available alternative.[17]

Citations to scientific articles can reflect learning on the part of industrial inventors through multiple channels. For instance, a firm may learn about a useful scientific discovery through an informal consulting relationship with an academic scientist or through the hiring of graduate students trained by that scientist rather than through a systematic and regular reading of the professional scientific literature. Even in these cases, the confluence of academic scientists' interest in rapid publication of significant discoveries combined with firms' legal obligation to cite relevant prior art means that citations to scientific articles will often show up in patent documents, providing a "paper trail" of knowledge diffusion, even when the particular means of knowledge diffusion was something other than the publication itself.

What my methodological approach clearly fails to measure is the contribution of "old science" to industrial invention. A significant component of the consulting work undertaken by university faculty consists of helping private industry understand and apply well-established – or,

---

[16] This result is also emphasized strongly in the interview-based evidence presented by Agrawal and Henderson [2002].

[17] Other recent studies using data on patent citations to scientific papers include work by Fleming and Sorenson [2000, 2001] and Lim [2001]. Neither of these studies focuses on the large change in citations to academic science over the course of the 1990s, which is the focus here.

"old" – scientific techniques and engineering principles, rather than helping firms incorporate the latest frontier science into their research agendas. Likewise, recent science and engineering graduates are often employed in functions that are quite far removed from the scientific frontier, but are nevertheless quite economically important to the financial success of their employers. This contribution will be completely missed by my approach. In such cases, there is no new patented invention incorporating recent science. But as the older literature on university-industry interaction has stressed, the propagation of "old" scientific and engineering knowledge to industry through training and consulting is a *long-standing feature of the American university system*. The new development stressed by the recent literature is the closer relationship between technology and relatively recent science. It is precisely this aspect of university-industry interaction that my methodological approach will most closely reflect.

III. Examining Patent Citations to Science: A Citations Function Approach

If I am to measure the relative importance of the four alternative hypotheses outlined in the introduction, then it is essential that I examine changes in patent citations to papers while controlling for growth and changes in the distribution across fields of the population of potentially cited papers, growth and changes in the distribution across fields of the population of potentially citing patents, and differences in the historical propensity for different categories of patents to cite science. While it would be impractical to do this for the universe of academic publications and U.S. patents, it has been possible for me to obtain and link the requisite data for the campuses and affiliated research units of the University of California, Stanford University, the California Institute of Technology (Caltech), and the University of Southern California. These are the principal sources of academic research in the state of California. Inference in this paper will be based on U.S. patent citations made to scientific articles generated by these institutions. There is no geographic restriction, however, on the location of the inventor of the citing patent.

The focus on California-based academic institutions as sources of science clearly limits the scope of this study, but it is also true that the geographic locus of innovative activity in the United States over the 1980s and, particularly, the 1990s, has shifted rather dramatically from the East Coast to California [Hicks et al. 2001]. One of the reasons given for this shift is the quality of the university science infrastructure in California, to which local firms are believed to have preferential access. Among other things, this paper will submit that belief to an empirical test.

Related research strongly suggests that the patterns in the citation data used in this study closely mirror national trends. In a companion paper [Branstetter 2003], I examine the *complete* set of nonpatent citations made by a random sample of 30,000 U.S. patents granted over the 1987-1997 period. I find that the distribution of patent citations to science across fields of science and technology in that random sample is very similar to that indicated in the current paper. This suggests that one of the key findings of the current paper – the concentration of patent citations to science in bioscence-related inventions – is not an artifact of my focus on California research universities. I also find a growth rate of patent citations to science in the random sample that is similar to that found in the raw data used in the current paper. Nevertheless, one must be sensitive to the potential difficulties involved in generalizing from my results to the entire American research university system. Wherever such difficulties arise, they are noted in the discussion of empirical results in sections IV and V.

From the University Science Indicators database generated by the Institute for Scientific Information, I have obtained comprehensive data on the publication of scientific articles by my sample of California research universities, by institution, year, and scientific field, from 1981-1997. These data are matched to data on patent citations made to these publications over the 1983-1999 (grant year) period, which are provided by CHI Research. CHI Research has developed a comprehensive data base of "non-patent references" made in U.S. patent documents. These references include citations to scientific journals, industrial standards, technical disclosures, engineering manuals, etc. The focus on this paper is on the subset of references made to articles appearing in peer-reviewed scientific journals. In the CHI Research database, references to scientific journals are put into a standardized format, and these data can then be matched to data on papers published in the more than 4,000 journals covered by the Science Citation Index (SCI).[18] Through this matching process, I obtain data on patent citations to peer-reviewed scientific articles generated by California research universities. To these data I match data on the universe of potentially citing U.S. utility patents granted over that same period, which is available from the NBER Patent Citation Database documented in Hall et. al. [2001].

Trends in scientific publications generated by California research universities for a subset of scientific disciplines are plotted in Figure III. Particularly strong growth can be observed in biomedical research, "physics" (an aggregate which includes materials sciences fields connected to semiconductors), and "engineering and technology."[19] Trends in U.S. patenting across different categories of technologies are similarly plotted in Figure IV. While patenting in all

[18] For a more detailed description of the database developed by CHI Research, see Narin et. al. [1997]. Further details are also available from the author upon request.

[19] Comparison of these data with similar data for all major American research universities shows that California academic publication closely mirrors national trends.

fields has increased over the sample period, particularly sharp increases can be seen in "drugs and medicine" and "computers and communications."[20]

The empirical framework I use for analyzing these data borrows from the work of Jaffe and Trajtenberg [1996, 2002]. In this framework, I model the probability that a particular patent, *p*, applied for in year *t*, will cite a particular article, *a*, published in year *T*. This probability is determined by the combination of an exponential process by which knowledge diffuses and a second exponential process by which knowledge becomes superceded by subsequent research.

This probability is referred to in the work of Jaffe and Trajtenberg [1996, 2002] as the *citation frequency*. It is a function of the attributes of the citing patent (P), the attributes of the cited article (a), and the time lag between them (t-T). It can be rendered in notation as

(1) $$p(a,P) = \alpha(a,P)\exp[-\beta_1(t-T)][1-\exp(-\beta_2(t-T))]$$

Attributes of the citing patent that I incorporate into my analysis include the application year, the technical field (based on the primary technology class assigned by the patent examiner), the type of entity owning the patent (based on the identity of the assignee), and the geographic location of the patent, based on the address of the inventor. Attributes of the cited article that I consider include the publication year, the scientific field of the article, and the institution with which the authors were affiliated at the time of publication.

Given these data, one could sort all potentially citing patents and all potentially cited articles into cells corresponding to the attributes of articles and patents. The expected value of the number of citations from a particular group of patents to a particular group of articles could be represented as

(2) $$E[c_{tcelTSL}] = (n_{TSL})(n_{tcel})\alpha_{tcelTSL}\exp[-(\beta_1)(t-T)][1-\exp(-\beta_2(t-T))]$$

where the dependent variable measures the number of citations made by patents in the appropriate categories of grant year (t), technology class (c), institutional type (e), and location of the citing patent's inventor (l) to articles in the appropriate categories of publication year (T), scientific field (S), and particular campus (L). The α's are multiplicative effects estimated relative to a benchmark or "base" group of patents and articles. In this model, unlike the linear case, the null hypothesis of no effect corresponds to parameter values of unity rather than zero. Equation (2) can easily be rewritten as

(3) $$\frac{E[c_{tcelTSL}]}{(n_{TSL})*(n_{tcel})} = \alpha_{tcelTSL}\exp[-\beta_1(t-T)][1-\exp(-\beta_2(t-T))]$$

[20] This graph does not break down patent trends by nationality of the inventor, but the fraction of patent grants awarded to domestic inventors has risen sharply in these two rapidly growing fields.

This is what Jaffe and Trajtenberg [1996] refer to as a *citations function*. If one adds an error term, then this equation can be estimated using nonlinear least squares. The estimating equation is thus

(4) $$p_{tcelTSL} = \alpha_t \alpha_c \alpha_e \alpha_l \alpha_T \alpha_S \alpha_L \exp[-\beta_1 (t-T)][1-\exp(-\beta_2 (t-T))] + \varepsilon_{tcelTSL}$$

where the dependent variable now measures the likelihood that a particular patent in the appropriate categories (grant year, technology class, institution type, and location) will cite an article in the appropriate categories (science field, source campus, and publication year).

Patents are placed into one of the following categories: computers and communications, chemicals, drugs and medicine, electronics, mechanical inventions, and a catch-all "other" category. These are the same categories for which patent growth is depicted in Figure III. Scientific articles are classified into the following fields: biology, biomedical research, chemistry, clinical medicine, engineering and technology, physics, and "other science." Patent assignees are classified into the following institutional types: public science institutions (predominantly universities, research hospitals, and government laboratories), firms, and other institutions. The division of patents on the basis of location of the inventor and the assignment of patents and papers into groups based on grant and publication year, respectively, are discussed below.

I estimate various versions of (4) using the nonlinear least squares estimation routine of the STATA software package. When doing so, I weight the observations by the square root of the product of potentially cited articles and potentially citing patents corresponding to the cell, that is

(5) $$w = \sqrt{(n_{tcel})*(n_{TSL})}$$

This weighting scheme should take care of possible heteroskedasticity, since the observations correspond to "grouped data," that is, each observation is an average (in the corresponding cell), computed by dividing the number of citations by $(n_{tcel})*(n_{TSL})$.

IV. Evidence from the Full Sample

Localization in Time and Geographic Space

Regression results from a version of (4) run on the full sample are given in Table I. Using the parameter values from this regression, it is also possible to graph out the double exponential function implied by our parameter estimates, giving us a sense of how the "citedness" of a particular group of articles by a particular group of patents changes over time. This is graphed out for our "base case" in Figure V. The base case in this regression corresponds to patents assigned to firms, where the first inventor resides in the U.S. outside the state of California. The base patent grant period is 1983-1987, and the base publication period is 1981-1985. The base science category is biology, the base patent category is chemistry, and the base institution is Stanford University.[21]

The shape of the curve graphically demonstrates the first key result of this section – namely that citations to academic science are, to some extent, localized in time. Citations to science appear almost immediately after article publication, and the citation function peaks at a lag of about eight years after article publication. These lags are measured here with respect to the grant date of the patent. An alternative specification measuring patents by *application* date finds a modal lag between publication and application of five to six years, implying fairly rapid spillovers of knowledge from science into industrial invention. While the estimated lag structure demonstrates that papers continue to receive some citations even at relatively long lags, the citation frequency declines steadily after the peak lag.

These results also provide evidence of concentration in geographic space. Citing patents are assigned to three categories based on the recorded addresses of the inventor: California inventors, U.S. inventors outside California, and non-U.S. inventors.[22] U.S. inventors outside California are the base category, so the coefficients imply that California-based inventors in a given technology class are nearly three times more likely to cite California academic science. Non-U.S. inventors are only about half as likely to cite California science as is the base category.

The intranational localization of knowledge spillovers implied by the California effect seems large. However, the current specification arguably does *not* control well for regional clustering of industrial R&D within the particular niches of the broad technology categories I have employed. A finer disaggregation of patent classes would likely attenuate the measured degree of localization. Furthermore, as can be seen in Figure VII, it is still the case that large numbers of citations are made by inventors far from California. In fact, one sees a "bicoastal"

---

[21] As commonly understood, biology is an aggregate that contains components closely associated with the bio nexus (molecular biology) and components that are arguably not closely connected to "biotech" (such as population ecology). In this paper, however, I have classified the subdisciplines of biology closely connected to the bio nexus as "biomedical research." Subdisciplines that remain within the biology aggregate used in this paper include such fields as ecology and "aquatic sciences." They are not closely connected to the bio nexus and, defined this way, "biology" would seem to be a reasonable base category. Note also that the institutional boundary of campuses like Stanford is drawn to include affiliated medical schools.

[22] The NBER Patent Citation Database only includes information on the address of the first inventor listed on the patent document, so that is the basis for geographical assignment of the patent undertaken here.

concentration of citations, reflecting the clustering of U.S. innovative activity in the Northeast and the West Coast.

Localization of Knowledge Flows in Technology Space and the "Changing Composition" Hypothesis

I find striking differences in the incidence of citation across fields of technology. Relative to the base category (chemicals), drug/medicine patents are 2.6 times more likely to cite science, whereas all other categories are substantially less likely to cite science. The typical patent in the least likely-to-cite category, mechanical patents, is only about 1% as likely to cite science as the typical chemical patent. The estimated gap between technology categories in citation propensity is quite substantial. Note that these estimated propensities control for the number of patents in these categories over time, so that these coefficients are properly interpreted as an estimate of the differential "per-patent" propensity to cite science.

Continuing in this theme, I can also allow different categories of science to display different propensities to be cited by patented technologies. Note that the citation function specification controls for the number of "citable papers" within these science categories over time, as well as the number of potentially citing patents across fields of technology, so the coefficients on science categories are akin to a "per-paper" measure of technological fertility. The coefficients in Table I suggest that a paper in the "biomedical research" field is *about 41 times* more likely to be cited in a patent than a paper in the base category of biology. Papers in "chemistry" and "clinical medicine" are nearly five times as likely to be cited as a biology paper, while papers in the other science categories are substantially less likely to be cited than biology papers.[23] The gap between the most and the least intensely citing technology categories is a factor of nearly two hundred.

As one can see in Figure IV, "biomedical" patenting has risen sharply over my sample period, both in absolute terms and relative to patenting in other technology categories. In fact, patenting in this area has risen more than four-fold. Likewise, as Figure III indicates, there has been substantial growth in publishing in bioscience areas by California research institutions. Even controlling for this growth, biotech patents are much more likely to cite science through the sample period, and bioscience papers are much more likely than other categories to be cited. This

[23] In results available upon request, I estimated an "academic production function" for the university systems studied in this section of the paper, in which the output measure was the count of publications generated in a scientific field by a particular campus in a particular year. This was regressed on measures of "inputs" to the research process. The results suggest that the higher "productivity" of the biomedical sciences is not driven purely by the increase in R&D funding in that field.

suggests that the aggregate trends in patent citations to science are driven largely by “biotech” patents citing “bioscience” papers. While there is growing citation activity outside this “bio nexus,” patent citations to science have, to date, been highly concentrated within it.

In another take on the “composition hypothesis,” I have also looked at patenting by different categories of assignees: firms, public science institutions (universities, research institutes, and research hospitals), and a grab-bag category of “other institutions” in the non-profit sector. Assignment of a patent to one of these categories is based on the typography of assignees developed in the NBER patent citation database. Relative to the base category of firms, public science institutions are nearly four times as likely to cite academic science, and “other institutions” are almost twice as likely to cite academic science, according to Table I. This is unsurprising, given the connection that is likely to exist between academic science and academic patenting. Because these institutional categories accounted for a small fraction of total U.S. patenting, even by the end of my sample period, it is still the case that the vast majority of patent citations to California academic science are made by the patents of industrial firms, not universities.[24]

Evidence on “Changes in Methods of Invention”

Having incorporated fixed effects associated with the citing field of technology, the cited field of science, the cited institution, and characteristics of the citing inventor/assignee, I can also make some inference about average changes in citation patterns over time across fields. Perhaps the most interesting finding here is that the propensity to cite academic science is evidently growing over time. This can be seen by examining the pattern of coefficients on the citing patent grant year cohort terms. They increase steadily from the “base category” of 1983-87.[25] Note that I have explicitly controlled for the fact that academic publications in the heavily cited branches of science have become more numerous and that there has been an increase in patenting in fields that heavily cite academic science. These results are consistent with the view that there has been *a change in the nature of invention* such that inventors now draw more heavily on academic science.

Evidence on Attorney-Driven Changes in Patent Citations to Scientific Papers

[24] This statement requires some qualification. University patenting is highly concentrated in a small number of fields. By the end of my sample period, university patenting accounted for roughly 15% of health care-related patenting. That being said, the overall results in Table I are robust to the removal of patents granted to “public science institutions” (primarily universities and research hospitals) from the sample. In fact, in some ways, they become even stronger. See Table III and the discussion on page 21.

[25] This pattern is quite robust to alternative aggregations of grant years into categories. Regression results demonstrating this are available from the author upon request.

These results could also be driven, at least in part, by an "attorney-driven" change in citation practice, and, in fact, interpretation of the measured increase in the per-patent propensity to cite academic papers is clouded by the issue of the so-called "spike patents."[26] In 1995, the effective period of monopoly granted to U.S. patent holders changed from 17 years after the grant date to 20 years from the filing date, in order to bring U.S. patent law more fully into compliance with the provisions of the TRIPs Agreement. This change took effect for patents filed after June 8, 1995. Patents filed prior to this deadline benefited from a "grandfather" provision – they were granted a monopoly of either 17 years from date of grant or 20 years from date of application, whichever was longer. Rejected patents re-filed after this deadline would also be subject to new evaluation criteria.

Applications submitted to the U.S. PTO more than doubled in May and June of 1995, and these applications, referred to as the "spike patents," carried an unusually large number of citations to science. This surge in patenting seems to have been driven in part by a rush to file in order to benefit from the "grandfather" timing provision. The increase in citations to science seems to have been driven in part by a desire to avoid having to refile rejected patents under the new rules. Applicants thus erred on the side of caution by making far more than the usual number of citations to scientific material. Patents applied for in this period were issued gradually over the next few years – dramatically increasing the average citations to science per patent in the overall data. Once the last of these applications was processed, average science citations per patent actually *fell*, as is illustrated in Figure VI. This kind of simple data tabulation might suggest that the connection between science and technology is weakening, after nearly a decade of rapid growth. That conclusion would be unwarranted, but it is likely that some of the movement in the aggregate data in the mid-to-late 1990s was "attorney-driven."

Within the context of my empirical approach, one potential remedy for this problem is to remove the spike patents from my data set and re-run the citations function. The results are shown in Table II, and it can be seen here (and in all subsequent tables, where the spike patents have been removed), that the basic qualitative features of the previous empirical results remain. In particular, the finding of an increase in per-patent propensity to cite scientific papers is robust to the removal of these patents.[27]

[26] This issue is also discussed in the 2002 issue of *Science and Engineering Indicators* and in Hicks et. al. [2001].

[27] Of course, removing the spike patents does not completely eliminate the possibility that measured changes in per-patent citation propensities reflect attorney-driven changes in citations practices. However, the desire the avoid litigation or impress examiners would presumably apply across different fields of technology. Likewise, the increasing availability of computerized databases, which reduce the costs of searching for scientific prior art, applies broadly across nearly all scientific fields. It is therefore striking

Evidence on Changes in Scientific Fertility

In the full sample, measures of per-paper "citedness" increase, relative to the base period, in the late 1980s and early 1990s, peaking in the 1989-92 period. They then seem to decline somewhat in the most recent period, but estimated per-paper citedness remains higher than in the base period. This fact would seem to provide reasonably strong evidence for the "changes in scientific fertility" hypothesis. However, this finding is *not* robust to the exclusion of university patents from the sample. The latter point is illustrated in Table III, which presents results based on a sample that excludes both spike patents and patents assigned to universities and to other "public science institutions," a category including research hospitals that often have links to universities. As can be seen, the apparent increase in per-paper citedness evaporates with this sample restriction. Other patterns in the results, however, are robust to this sample restriction. The measured localization of spillovers within the bio nexus remains after dropping university patents, and the measured increase over time in per-patent propensity to cite science becomes more pronounced.

Summarizing the Lessons from the Full Sample

Once we exclude "spike patents," it seems that trends in the data are best explained by a combination of the "changing composition" story and the "changing methods of invention" story. However, one needs to put the relative importance of these issues into perspective. To that end, it is useful to examine Table IV, which presents results from a series of hypothesis tests. It is certainly true that the data reject the imposition of the constraint that methods of invention have not changed – or, more precisely, that per-patent propensities to cite science have not changed broadly across fields of technology. The value of the Wald test associated with this parameter restriction (see the second column, third row) is 1,256.6, and this easily exceeds the critical value of the Chi-Square distribution at the appropriate degrees of freedom. But the degradation in model fit generated by this constraint is small. Relative to the unrestricted model, the adjusted R-squared of the restricted model declines by only about 1%. This can be inferred by comparing

---

that patent citations to science are so tightly concentrated in that narrow "nexus" of sciences and technologies where the recent literature suggests that the intellectual interaction is the strongest. Furthermore, conversations with patent attorneys indicate that, while patent attorneys and patent examiners often insert citations to previous patents unknown to the inventor into patent applications for legal, strategic, or procedural reasons, they are much less likely to insert citations to the academic literature, largely because they are much less familiar with it than is the inventor. In other words, citations to science are likely to be a purer measure of "knowledge spillovers" than are patent citations to patents. See Jaffe, Fogarty, and Banks [1998].

values in the third column – the adjusted R-squared values associated with the restricted models – with the adjusted R-squared of the unrestricted model given on the next to last row.

In striking contrast, imposing the constraint that the relative propensity of different patent classes to cite science is the same causes the adjusted R-squared to fall by about 67%, and imposing this constraint and the constraint that the relative citedness of different categories of science is the same causes the adjusted R-squared to fall by about 85%, relative to the unrestricted model. In other words, changes in the distribution of patenting across technologies and changes in the distribution of publications across fields explain much more of the total variance in patent citations to science than does average changes across fields in per-patent citation behavior over time. Now, it is possible – even likely – that part of the substantial expansion in biotech patenting has been driven by increasing knowledge spillovers from university science. This is an idea that will be probed more thoroughly in the next section. Nevertheless, the scope of these increased knowledge spillovers from university science seems to be rather narrowly confined to a small set of technologies.

The evidence in Tables I-III comes from a version of the citation function in which I constrain the obsolescence parameter to be the same across categories of technology. Following Jaffe and Trajtenberg [2002], I can allow this parameter to differ across patent technology categories. Results from such a regression are omitted for reasons of space. Allowing this parameter to vary does not change the qualitative patterns in the other results. Not surprisingly, estimated obsolescence is significantly faster for computers/communications and electronics, in the sense that the differences are both qualitatively large and statistically significant.[28]

V. Evidence from Data Subsamples – the Bio Nexus vs. the IT Nexus

Partitioning the Data

For every year of my sample period, roughly 70-90% of the total citations made by patents to scientific articles are made within the bio nexus. Given the extent to which aggregate numbers of citations are driven by biotech, I break the data into a biotech-only subsample and a subsample from which papers in bioscience and patents in biotechnology are excluded.[29] This partition of the data allows me, at least in principle, to examine changes in the bioscience-biotechnology nexus in some detail. Then, I can separately estimate the key parameters of the

---

[28] These results are available from the author upon request.

[29] The discipline of chemistry is somewhat unique in that it includes subdisciplines that are closely connected to the bio nexus and other subdisciplines that are completely unrelated. Given this dual nature, I include chemical patents and chemistry papers in both subsamples.

citations function for the "non-bio" subsample, such that the parameters of diffusion, geographic concentration, etc. – constrained to be the same across fields of science and technology -- are not driven by observations in the bio nexus. The pattern of knowledge diffusion from science to invention may be quite different outside the bio nexus; partitioning the data in this way allows me to quantify those differences.[30]

Evidence from the Bio Nexus

Evidence from the bio nexus is presented in Table V. Table V maintains the same aggregation scheme across patent classes as Tables I-III, but uses only data from the bio nexus in estimating the citations function. The "biomedical research" cluster of scientific fields is broken up into the "core fields" of biochemistry, biophysics, and molecular biology on the one hand and the remaining fields of biomedical research on the other. This sample excludes spike patents, but includes patents assigned to universities, research institutes, and research hospitals.

The qualitative results are similar to those estimated in the full sample. In particular, one finds, even within the bio nexus, statistically significant evidence of an increase in the per-patent propensity to cite science over time. In other words, even within this nexus, where citation activity has always been strongest and where the number of patents has been growing rapidly, the connection between industrial research and academic science seems to have grown substantially over time. The estimates on the grant year coefficients suggest that per-patent citation propensities had increased by nearly 80% (relative to the base period) by 1997-99. In this particular sample of the data, the measured increase in per-paper citedness – our measure of changes in scientific fertility – remains roughly what it was in the overall sample including public science patents. That is, it suggests an increase, albeit non-monotonic, in scientific fertility over time. While this result does not survive the exclusion of public science patents, the result of an increase in per-patent citation propensity does. Looking carefully within the bio nexus itself, one finds strong evidence of a change in the method of invention over my sample period.

The finding of an increase over time in the per-patent intensity to cite science is supported by a number of studies of the pharmaceutical and biotech industries. From its inception, the biotechnology industry has been closely linked to university-based science.[31] Furthermore, over the course of the 1980s and 1990s, established pharmaceutical companies have

---

[30] A potential downside to this partition is that we lose "cross-nexus" citations, such as citations made by biotech patents to papers in mathematics and computer science. It is true that we observe an increase in such cross-nexus citations over time, probably reflecting the increasing importance of such fields as "bioinformatics," but the aggregate numbers of these cross-nexus citations remain small, even in the most recent periods.

[31] See Kaplan and Murray [2003], Kenney [1986], and Gambardella [1995].

increasingly drawn on recent scientific developments in their efforts to improve the efficiency of their drug discovery programs.[32] While the received literature has not yet tried to *quantify* the changes in this intensity of industrial borrowing from academic science over time, the basic trends in my data are reasonably consistent with the qualitative descriptions of changes over time in the existing literature. Because of limitations in the time dimension of my sample, I am picking up relatively little of the impact of genomics, proteinomics, and bioinformatics on the most recent cohorts of biotech invention. While the ultimate impact of these relatively new disciplines on industrial invention remains to be seen, it is certainly possible that a future update of my regressions may find *further* acceleration of the shift in industrial invention toward greater reliance on academic science.

Have these spillovers from academic science actually raised inventive productivity in the bio nexus? A casual examination of the aggregate evidence would suggest an affirmative answer. According to NSF data, total real R&D outlays from both public and private sources associated with the life sciences nearly doubled between 1985 and 1995. However, U.S. patenting in the bio nexus more than tripled over this period, which would seem to imply a considerable increase in R&D productivity for the nexus as a whole. On the other hand, studies by others have suggested that the patent to real R&D ratio has fallen substantially for large U.S.-based pharmaceutical companies – an important component of the nexus – over much of my sample period.[33]

This discrepancy calls for careful empirical analysis of the relationship at the firm level between the incorporation of academic science into industrial R&D and its effect on research productivity. Within the bio nexus, scientific developments have apparently provided new technological opportunities for private-sector researchers to explore. The observed increase in R&D expenditures, the increase in patenting, and the striking increase in patent citations all seem to bear witness to this. It is not yet clear, however, that the new domains opened up by academic science will prove to be more fertile than the domains that preceded them. This is a point to which I will return in the concluding section and remains a focus of current research.

Evidence from the IT Nexus

The non-biotech subsample generates a significantly different pattern of results. The aggregate patent classes used are computers and communications (IT), chemistry, general electronics, mechanical inventions, and a catch-all "other" category. Science aggregates are

[32] See, among others, Cockburn, Henderson, and Stern [1999], who provide a useful qualitative description of these changes, then go on to document their implications for relative firm performance over time within the pharmaceutical industry.
[33] See, among others, Henderson and Cockburn [1996].

engineering and technology, chemistry, physics, and a catch-all "other science" category. The other categories remain as before. Note that we are estimating roughly the same number of parameters for our non-biotech subsample as in the full sample, even though we have only a small fraction of the number of observations of patent-article citations. The relative thinness of the data here means that our parameter estimates need to be regarded with an extra measure of caution, even when they are statistically significant according to the conventional thresholds. Results are given in Table VI.

Patent-article citation activity outside the bio nexus is clearly concentrated in a secondary "IT" nexus. The IT patent classes cite science most frequently, displaying a propensity to cite that is nearly 18 times as high as the base category of chemistry. General electronics patents are more than 7 times as likely to cite science, while mechanical patents are three times as likely. Articles in the physics fields are more than 44 times more likely to be cited than base category (chemistry) articles. The physics aggregate includes fields that relate to semiconductors and advanced materials. The engineering/technology aggregate (which includes computer science) is the next most highly cited, with a citedness per paper that is about 8 times greater than the base category. The rest of the sciences are significantly less likely to be cited. Incidentally, these results suggest that much of the citation activity involving chemical patents and chemistry papers comes from the bio nexus. Once chemistry is separated from those technologies and disciplines, it ceases to stand out in terms of patent-article citation activity.

In a striking contrast with earlier results, geographic localization seems to be much higher in this subsample. California-based inventors display a much higher likelihood of citing California science than the base (non-California U.S.) category of inventors. While intra-national localization is much higher, *international* localization is lower – the tendency of non-American inventors to cite California science is nearly 75% as high as that of non-California American inventors, and the difference between them and the base category is not statistically significant. This pattern of results could very well reflect the increasing geographic concentration of the U.S. information technology industry in California, as well as strong growth by inventors based outside the United States (particularly East Asia) in patenting in IT-related classes.

Another contrast with earlier results is a much higher propensity (relative to industrial firms) for patents generated by public science institutions to cite science. Public science institutions are more than 38 times as likely to cite science as are firm patents, controlling for patent category. Patents held by "other institutions," are less likely than firms to cite science in these fields, corresponding to the less significant role played by this category of assignee in non-biotech patenting.[34] The pattern of campus coefficients also highlights the unique role played by Stanford University within the sample. While, within the full sample and the bio nexus, a number of other institution's "campus effects" were nearly as high, or even higher, than Stanford's, in the non-bio subsample no other institution comes remotely close to Stanford's implied relative level of academic fertility. As with the estimates of geographic localization, it seems these data reflect the "Silicon Valley" phenomenon.

The patterns suggested by the coefficients on patent grant year cohorts and paper publication year cohorts also differ from those in previous regressions. Controlling for changes in the volume and distribution of publications and patents, all periods display a substantially greater per-patent propensity to cite science than the base period. Rather than the monotonic increase one saw in the full sample, the pattern here looks more like a step function, with a sharp increase in the late 1980s. Although the increase relative to the base period is higher than in the bio nexus, one has to keep in mind that the absolute numbers of citations in this category remains *much* smaller than in the bio nexus. The increase in per patent propensity to cite science, combined with a sharp increase in patenting in the IT-related classes, explains most of the aggregate increase in citations to science outside the bio nexus.

While much of the recent qualitative literature on university-industry interaction has focused on the extensive borrowing from science taking place in the bio nexus, this activity is less well documented outside that nexus.[35] Nevertheless, the timing of the increase in per-patent propensity to cite science noted above is roughly coincident with two major changes in patenting – a substantial increase in patenting by semiconductor firms, especially the so-called "fabless" IC design firms, and a sharp increase in software patenting. The semiconductor industry has always had strong links to science [Hicks et al., 2001, Lim, 2003], but, as Hall and Ziedonis [2001] have showed, firms in this industry began sharply increasing their patenting in the 1980s and 1990s. Furthermore, new entrants (the so-called "fabless" design firms) emerged that were often closely linked to university engineering departments.

The increase in software patenting followed changes in U.S. patenting law and practice which expanded the ability of software inventors to patent, rather than copyright, their inventions [Bessen and Hunt, 2003]. There was little patent "prior art" to cite, so patents in this area have tended to cite more nonpatent prior art, including the relevant academic work in computer science

---

[34] However, dropping "public science institution" patents from the sample does not qualitatively change the nature of the other results.

[35] Arora and Gambardella [1994] present many examples of what might be called a more "scientific" approach to industrial research outside the bio nexus, and stress the centrality of information technology in driving this shift.

and related fields. Software patents can be difficult to track, and the exacting timing of the measured increase depends on ones definition of a software patent. Nevertheless, some observers suggest that there was a sharp increase in software patenting at the end of the 1980s.[36]

The final result to note from our exploration of citation activity outside the bio nexus is that recent cohorts of papers are not more likely to be cited. In fact, the per-paper citedness measures have sharply plummeted, even when one includes public science patents in the sample. One can also see that the estimated obsolescence coefficient is *substantially* higher than the overall sample, while the diffusion parameter is lower. These results need to be viewed together. On average, the gap between paper publication and patent citation is *much* shorter than it is in the bio nexus, such that very recent science is much more likely to get cited. Controlling for this short gap, however, there is no evidence that the most recent cohorts of papers generate more knowledge spillovers. In fact, the estimated decline in per-paper citedness is so sharp that the substantial increase in publications in these disciplines fails to make a positive contribution to total citations. In general, it seems that citations to science in these categories arrive more quickly, decay more rapidly, and peak at a lower level.

Framing these results in light of the alternative hypotheses stated in the introduction, it seems clear that the increase in citations outside the bio nexus has been driven almost entirely by composition effects – both in terms of fields of technology, fields of science, and institutions -- and "changing methods of invention." In that sense, results here are broadly consistent with those discussed earlier. However, it must be stressed that citation activity in the secondary IT nexus identified in these data is substantially lower than that within the bio nexus – so much so that the IT nexus does not even show up in the full sample. The explosion of IT patenting in recent years has been even more dramatic than that of bio nexus patenting, but the relative paucity of citations to science among these patents suggests that knowledge spillovers from academia have almost certainly *not* played the primary role in generating this patent explosion.

VI. Conclusions and Extensions

What is driving the remarkable increase over the last decade in the propensity of patents to cite academic science? Does this trend indicate that stronger knowledge spillovers from academia have helped drive the surge in innovative activity in the U.S. in the 1990s? This paper has sought to shed light on these questions by using a common empirical framework to assess the

[36] See Bessen and Hunt [2003] who discuss the problems involved in measuring software patenting and provide alternative counts of software patents over time. Some of these series increase quite sharply in the late 1980s.

relative importance of various alternative hypotheses in explaining the growth in patent citations to science. My analysis supports the notion that the nature of U.S. inventive activity has changed over the sample period, with an increased emphasis on the use of the knowledge generated by university-based scientists in later years. That being said, knowledge flows from academia to industry, as they are measured in this paper, have been overwhelmingly concentrated in the bio nexus throughout the sample period. While scientific breakthroughs generated by academic researchers, particularly in the life sciences, have generated new "technological opportunities," these new opportunities are evidently limited in scope. Stronger knowledge spillovers from academia provide, at best, only a partial explanation of the recent surge in industrial innovation evident in U.S. patent statistics.

In my introduction, I laid out four alternative hypotheses that could possibly explain the sharp increase over time in the number of patent citations to science: the "increasing scientific fertility hypothesis," the "attorney-driven" hypothesis, the "changing method of invention hypothesis," and the "changing composition of invention" hypothesis. When one excludes "public science" patents, there is little robust evidence that the scientific fertility of more recent cohorts of papers is increasing over time. There is ample evidence of dramatic "attorney-driven" fluctuations in the overall level of citation, but these can be largely linked to a one-time change in U.S. patent law in the mid-1990s, and controlling for these "attorney-driven" changes in citations practice does not affect the qualitative importance of the "changing composition" and "changing methods of invention" hypotheses in explaining overall trends in the data. It is the combination of these latter two hypotheses that are most consistent with the data.

My results also speak, albeit indirectly, to the mechanisms by which knowledge flows from universities to industry. The patenting and subsequent commercial licensing of university-generated knowledge does *not* seem to be a necessary condition for useful knowledge flows to take place. In fact, the general tenor of my results is unaltered by the removal of all patents assigned to "public science institutions" from the database. This seems to result from the fact that, even today, only a tiny fraction of the commercially relevant science generated by universities is ever patented or licensed. It does not follow from this that the recent public policy focus, in the United States and elsewhere, on the establishment of technology licensing offices in universities and the encouragement of licensing activity is necessarily misplaced. Nevertheless, the results contained herein suggest that university patents and associated licenses are neither the only nor the most important means by which university-generated knowledge is transferred to industry.

The findings of this paper also suggest some interesting avenues for future research. As I have already noted, at the current level of aggregation used in this paper, it is difficult to come to a definite conclusion about the impact of knowledge flows from academia on the research productivity of citing firms. By tracking the patents and research inputs of individual firms over time, I would be able to bring to bear all of the usual panel data econometric techniques. These could provide useful leverage in determining whether increases in the intensity of a firm's citation of academic science actually lead to substantially higher levels of innovative productivity.[37] At the firm level, I can also go much further in terms of exploring how different firms have differentially benefited from university-based scientific research. The conventional wisdom suggests that university-linked start-up firms have been an important vehicle for the mediation of knowledge flows from universities to industry. With rich, firm-specific, time-varying data on the characteristics of citing firms, one could explore these and related hypotheses much more thoroughly. Pursuing such analyses at the firm level is the focus of current research.

The results of the paper also suggest a class of theoretical models that could guide further empirical analyses. Nearly twenty years ago, Evenson and Kislev [1976] proposed a model of the interaction between basic science and applied research. I have already argued that trends in patenting and patent citations within the bio nexus seem broadly consistent with their concept of a basic scientific breakthrough opening up new "search distributions" for applied research. Further theoretical work, building on the work of Evenson and Kislev and the subsequent work they inspired, such as Kortum [1997], may prove to be a useful complement to the empirical work described above.

[37] Preliminary empirical research by the author suggests that this is indeed the case. In keeping with the pattern of results in the current paper, the positive impact on research productivity appears to be substantially higher in the bio nexus than elsewhere.

Bibliography

Adams, James, 1990, "Fundamental Stocks of Knowledge and Productivity Growth," *Journal of Political Economy* 98: 673-702.

Adams, James and Zvi Griliches, 1996, "Research Productivity in a System of Universities," NBER working paper no. 5833.

Agrawal, Ajay, 2001, "Research on University-to-Industry Knowledge Transfer: Framework of Existing Literature and Future Questions," *International Journal of Management Reviews*.

Agrawal, Ajay and Iain Cockburn, 2003, "The Anchor Tenant Hypotheses: Exploring the Role of Large, Local, R&D-Intensive Firms in Regional Innovation Systems," forthcoming in the *International Journal of Industrial Organization*.

Agrawal, Ajay and Rebecca Henderson, 2002, "Putting Patents in Context: Exploring Knowledge Transfer from MIT," *Management Science*, Vol. 48, No. 1.

Arora, Ashish and Alfonso Gambardella, 1990, "Complementarity and External Linkages: the Strategies of the Large Firms in Biotechnology," *Journal of Industrial Economics*, Vol. 38, pp. 361-379.

Arora, Ashish and Alfonso Gambardella, 1994, "The Changing Technology of Technological Change: General and Abstract Knowledge and the Division of Innovative Labor," *Research Policy*, Vol. 23, pp. 523-532.

Arora, Ashish, Andrea Fosfuri, and Alfonso Gambardella, 2001, *Markets for Technology: The Economics of Innovation and Corporate Strategy*, MIT Press.

Audretsch, David and Paula Stephan, 1996, "Company-Scientist Locational Links: The Case of Biotechnology," *American Economic Review*, Vol. 86, No. 3.

Barnes, Michael, David Mowery, Arvids Ziedonis, 1998, "The Geographic Reach of Market and Nonmarket Channels of Technology Transfer: Comparing Citations and Licenses of University Patents," working paper.

Bessen, James and Robert Hunt, 2003, "An Empirical Look at Software Patents," unpublished working paper.

Branstetter, Lee, 2003, "Exploring the Link Between Academic Science and Industrial Innovation," unpublished working paper.

Cockburn, Iain and Rebecca Henderson, 2000, "Publicly Funded Science and the Productivity of the Pharmaceutical Industry," paper prepared for the NBER Conference on Science and Public Policy.

Cockburn, Iain, Rebecca Henderson, and Scott Stern, 1999, "The Diffusion of Science Driven Drug Discovery: Organizational Change in Pharmaceutical Research," NBER Working Paper No. 7359.

Cockburn, Iain and Rebecca Henderson, 1998, "The Organization of Research in Drug Discovery,"*Journal of Industrial Economics*, Vol XLVI, No. 2.

Cohen, Wesley, Richard Florida, Louis Randazzese, and John Walsh, 1998, "Industry and the Academy: Uneasy Partners in the Cause of Technological Advance," in R. Noll, ed., *Challenges to the Research University*. Washington, D.C.: Brookings Institution

Cohen, Wesley and Daniel Levinthal, 1989, "Innovation and Learning: The Two Faces of R&D," *Economic Journal*, no. 99, pp. 569-596.

Darby, Michael and Lynne Zucker, 2003, "Grilichesian Breakthroughs: Inventions of Methods of Inventing and Firm Entry in Nanotechnology," NBER working paper. 9825.

DeLong, J. Bradford, 2001, "Do We Have a 'New' Macroeconomy?," *Innovation Policy and the Economy*, Vol. 2, MIT Press.

Evenson, Robert and Yoav Kislev, 1976, "A Stochastic Model of Applied Research," *Journal of Political Economy*, Vol. 84, pp. 265-282.

Faulkner, Wendy and Jennifer Senker, 1995, *Knowledge Frontiers: Public Sector Research and Industrial Innovation in Biotechnology, Engineering Ceramics, and Parallel Computing*, Oxford: Clarendon Press.

Fleming, Lee and Olav Sorenson, 2001, "Science as a Map in Technological Search," working paper.

Gambardella, Alfonso, 1995, *Science and Innovation: The U.S. Pharmaceutical Industry during the 1980s*, Cambridge: Cambridge University Press.

Gordon, Robert J., 2000, "Does the 'New Economy' Measure up to the Great Inventions of the Past?" NBER Working Paper no. 7833.

Gordon, Robert J., 2002, "Technology and Economic Performance in the American Economy," NBER Working Paper no. 8771.

Henderson, Rebecca and Iain Cockburn, 1996, "Scale, Scope, and Spillovers: the Determinants of Research Productivity in Drug Discovery," *Rand Journal of Economics* 27, pp. 32-59.

Henderson, Rebecca, Adam B. Jaffe, and Manuel Trajtenberg, 1998, "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988," *Review of Economics and Statistics*, 119-127.

Hicks, Diana, Tony Breitzman, Dominic Olivastro, and Kim Hamilton, 2001, "The Changing Composition of Innovative Activity in the US – A Portrait Based on Patent Analysis," *Research Policy* 30, pp. 681-703.

Jaffe, Adam, 1989, "The Real Effects of Academic Research," *American Economic Review*, 79 [5], pp. 957-70

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson, 1993, "Geographic Localization of

Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, Vol. CVIII, No. 3.

Jaffe, Adam and Manuel Trajtenberg, 2002, *Patents, Citations, and Innovations: A Window on the Knowledge Economy*, MIT Press.

Jaffe, Adam and Manuel Trajtenberg, 1996, "Flows of Knowledge from Universities and Federal Labs: Modeling the Flow of Patent Citations over Time and across Institutional and Geographic Boundaries," NBER working paper no. 5712.

Jaffe, Adam, Michael Fogarty, and Bruce Banks, [1998], "Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation," *Journal of Industrial Economics*, Vol. XLVI, No. 2.

Jensen, Richard and Marie Thursby, 1999, "Proofs and Prototypes for Sale: The Licensing of University Inventions," *American Economic Review*, 91, pp. 240-259.

Kenney, Martin, 1986, *Biotechnology: The University-Industrial Complex*, New Haven: Yale University Press.

Kortum, Samuel, 1997, "Research, Patenting, and Technological Change," *Econometrica*, Vol. 65, pp. 1389-1419.

Kortum, Samuel and Joshua Lerner, 1998, "Stronger Protection or Technological Revolution: Which is Behind the Recent Surge in Patenting?" *Carnegie-Rochester Conference Series on Public Policy*, 48, pp. 247-304.

Kortum, Samuel and Joshua Lerner, 2000, "Assessing the Contribution of Venture Capital to Innovation," *Rand Journal of Economics*, 31, pp. 674-692.

Kortum, Samuel and Joshua Lerner, 2003, "Unraveling the Patent Paradox," unpublished working paper.

Lach, Saul and Mark Schankerman, 2003, "Incentives and Invention in Universities," unpublished working paper.

Lieberman, Marvin, 1978, "A Literature Citation Study of Science-Technology Coupling in Electronics," *Proceedings of the IEEE*, 36 [1], pp. 5-13.

Lim, Kwanghui, 2001, "The Relationship between Research and Innovation in the Semiconductor and Pharmaceutical Industries," working paper.

Mansfield, Edwin, 1995, "Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing," *The Review of Economics and Statistics* 77: 55-65.

Marquis, D. G. and T. Allen, 1966, "Communication patterns in Applied Technology," *American Psychologist*, 21, pp. 1052-1060.

Mowery, David, Richard Nelson, Bhavin Sampat, and Arivds Ziedonis, 1998, "The Effects of the

Bayh-Dole Act on U.S. University Research and Technology Transfer: An Analysis of Data from Columbia University, the University of California, and Stanford University," working paper

Murray, Fiona and Sarah Kaplan, 2001, "The Role of Social Evidence in Shaping a Technological Discontinuity: The Evolution of Biotechnology," unpublished working paper.

Narin, Francis, Kimberly Hamilton, and Dominic Olivastro, 1997, "The Increasing Linkage Between U.S. Technology and Public Science," *Research Policy* 197: 101-121.

National Science Foundation, 2002, *National Science and Engineering Indicators*.

Office of Technology Transfer, University of California, 1997, Annual Report: University of California Technology Transfer Program. Oakland, CA: University of California.

Price, Derek, 1965, "Is Technology Historically Independent of Science? A Study in Statistical Historiography," *Technology and Culture*, 6, pp. 553-568.

Rosenberg, Nathan and Richard Nelson, 1994, "American Universities and Technical Advance in Industry," *Research Policy*, 23, pp. 323-348.

Schmookler, Jacob, 1966, *Invention and Economic Growth*, Cambridge, MA: Harvard University Press.

Shane, Scott, 2000, "Prior Knowledge and the Discovery of Entrepreneurial Opportunities," *Organization Science*, 11, pp. 448-469.

Shane, Scott, 2001, "Technological Opportunities and New Firm Creation," *Management Science*, 47, pp. 205-220.

Sherwin, C. W. and R. S. Isenson, 1967, "Project Hindsight," *Science*, 156, pp. 1571-1577.

Sorenson, Olav and Lee Fleming, 2001, "Science and the Diffusion of Knowledge," working paper.

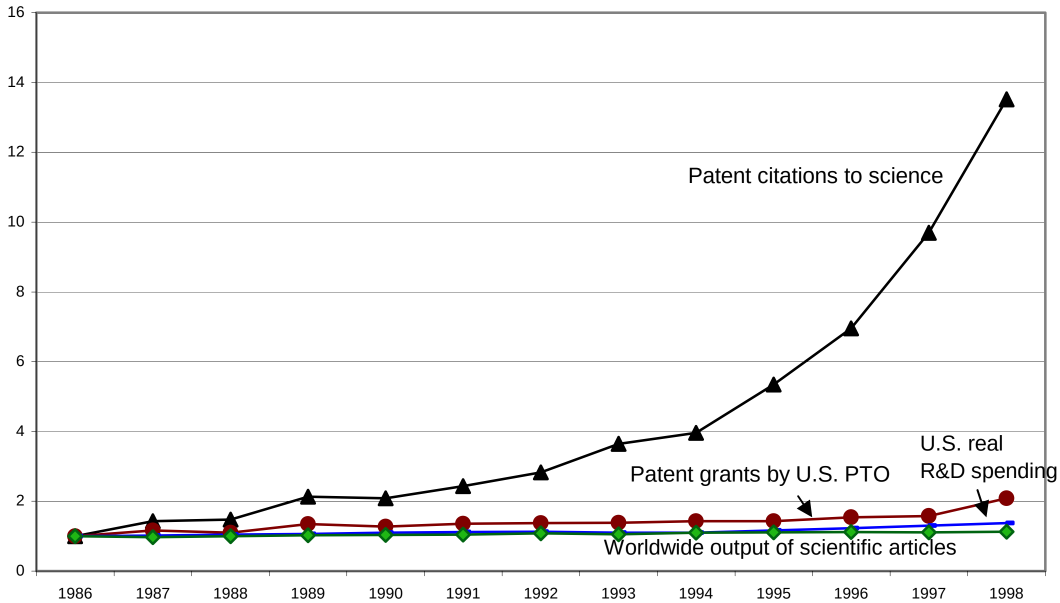
Stephan, Paula, 1996, "The Economics of Science," *Journal of Economic Literature* 34: 1199-1235.

Stern, Scott, 1999, "Do Scientists Pay to Be Scientists?" NBER Working Paper No. 7410.

Thursby, Jerry and Marie Thursby, 2002, "Who is Selling the Ivory Tower? Sources of Growth in University Licensing," *Management Science*, 48, pp. 90-104.

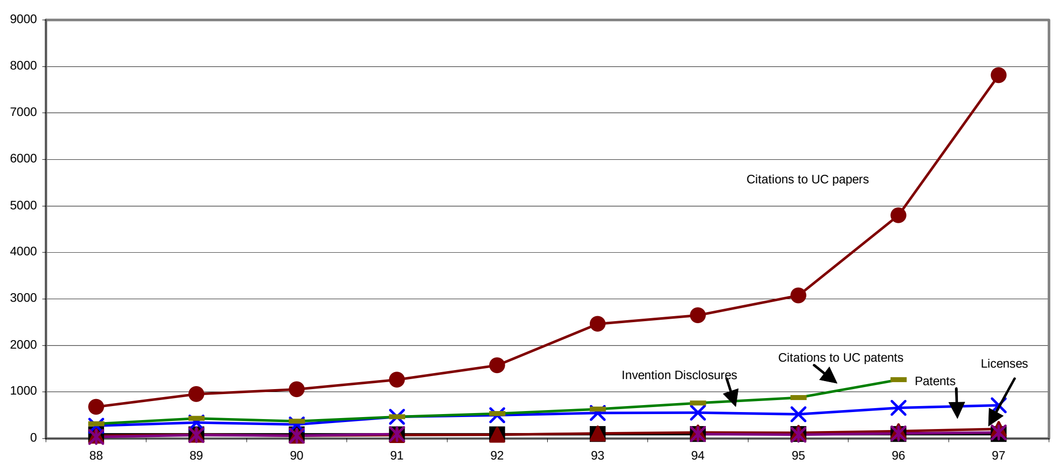
Zucker, Lynne, Michael Darby, and Michael Brewer, 1998, "Intellectual Capital and the Birth of U.S. Biotechnology Enterprises," *American Economic Review*, 88: 290-306.

Figure I Patent Citations to Academic Science
Series are scaled so that 1986 values are equal to 1



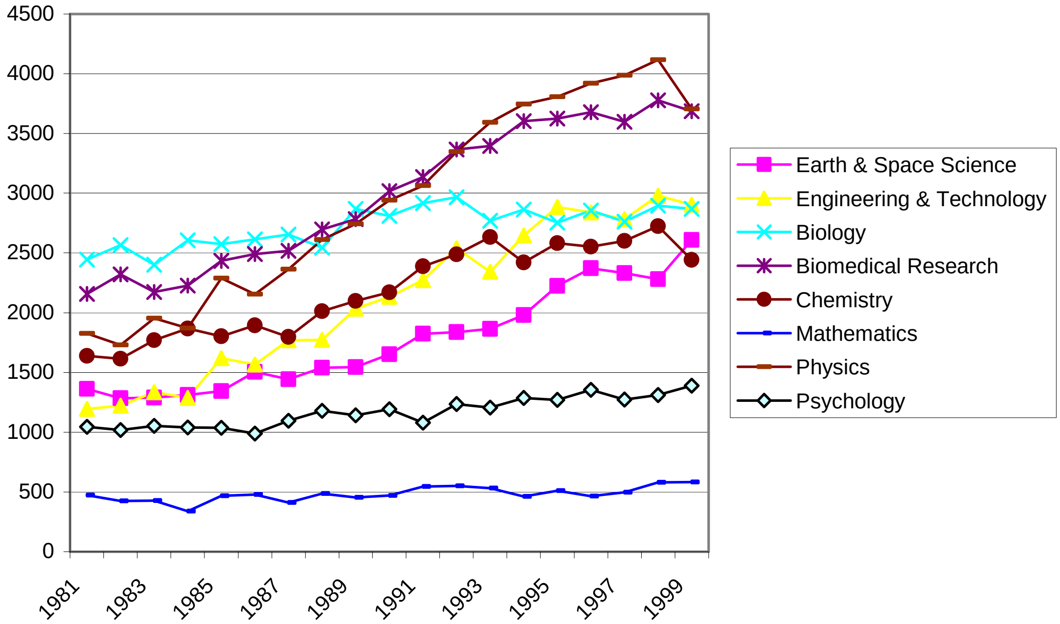
Source: National Science and Engineering Indicators, 2002

Figure II Citations to UC papers vs other indicators



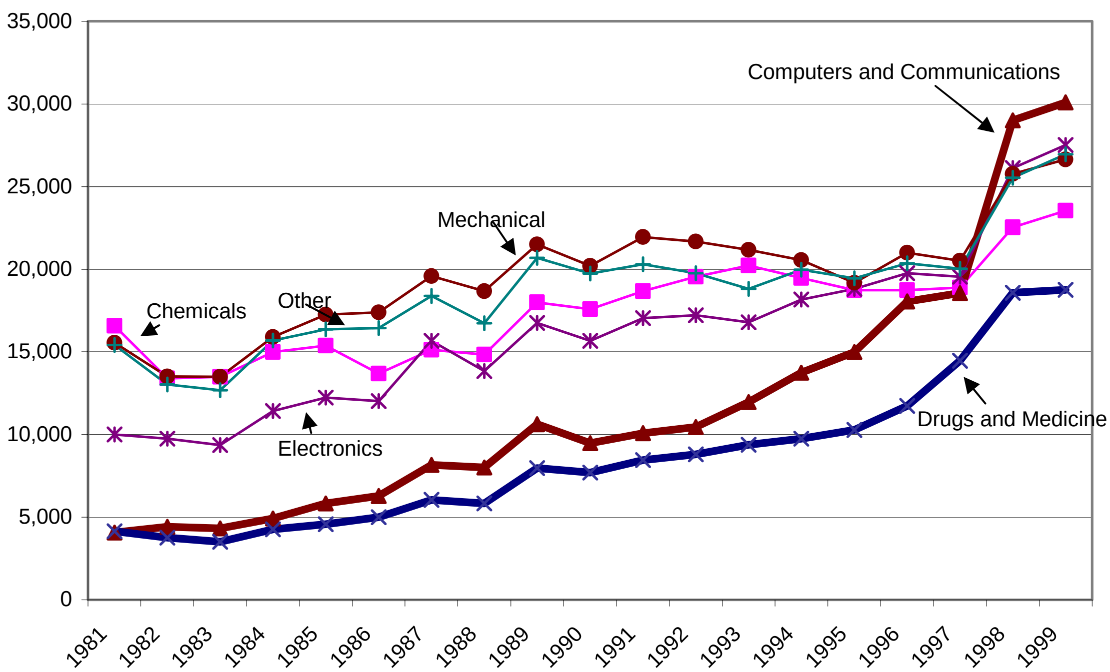
Source: Author's calculations based on data from the University of California Technology Transfer Office annual report, AUTMN, the NBER Patent Citation Data Base, and CHI Research.

Figure III
Growth in California Academic Publishing, Excluding Clinical Medicine



Source: University Science Indicators, 1999

Figure IV Patent Grants by Technology Category, 1981-1999

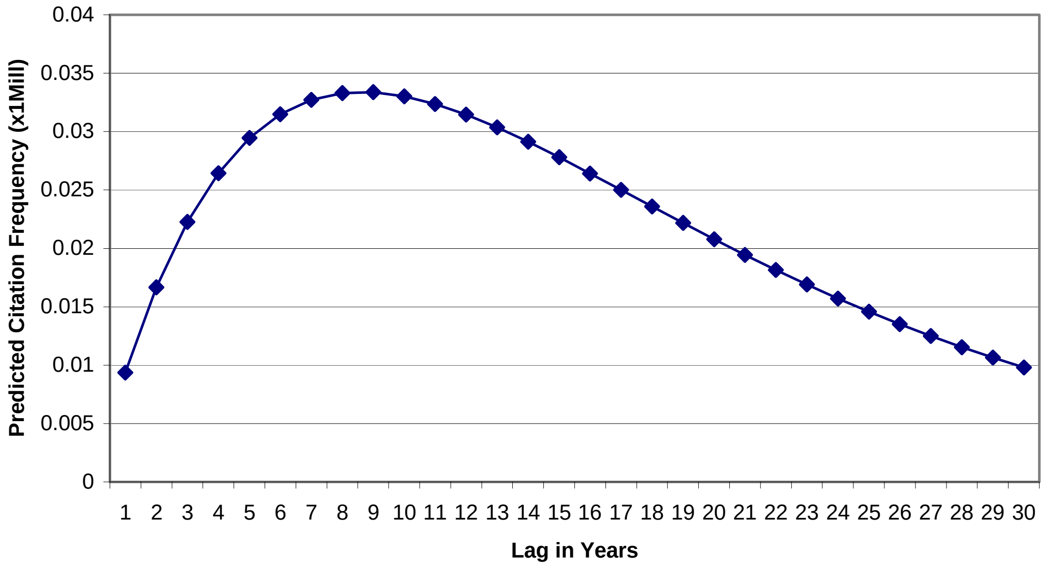


Source: NBER Patent Citation Database

Table I   Citation Function Results, Full Sample

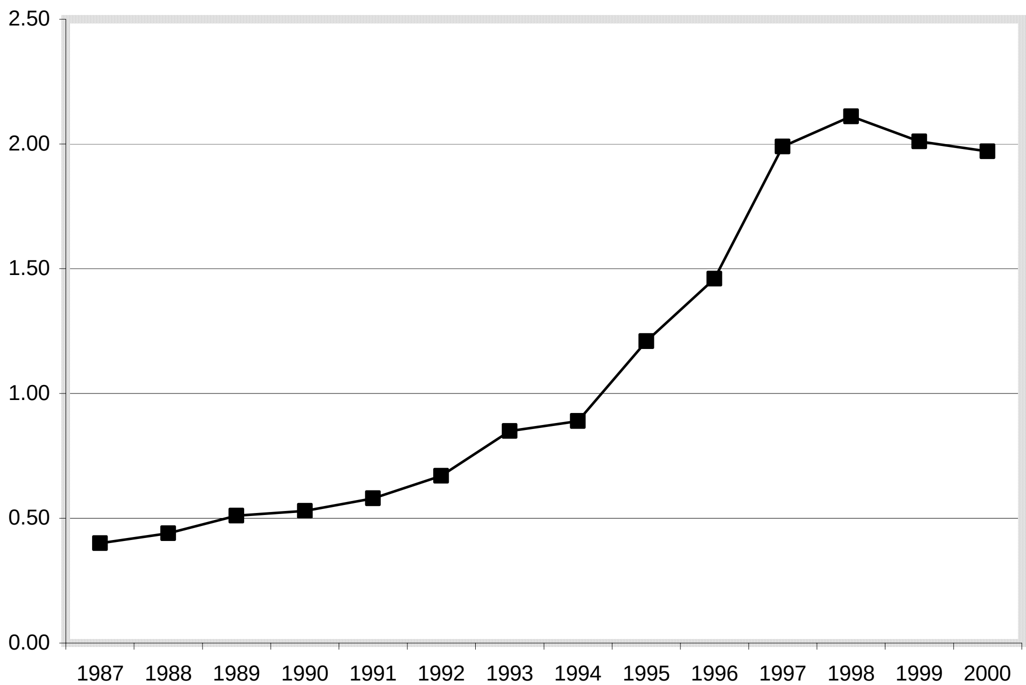
Variable	Coefficient	T-statistic for H_0: Parameter=1
Computers and Communications	0.04	-150.72
Drugs/medicine	2.44	91.93
Electronics	0.05	-157.91
Mechanical	0.01	-132.68
Other	0.05	-119.99
Biomedical research	41.02	10.11
Chemistry	4.75	8.03
Clinical Medicine	5.35	8.36
Eng/Technology	0.25	-7.94
Other Science	0.37	-7.00
Physics	0.49	-5.31
Caltech	1.19	23.93
Berkeley	0.57	-82.34
Davis	0.42	-112.52
Irvine	0.44	-93.81
Los Angeles	0.39	-128.68
Riverside	0.26	-110.96
Santa Barbara	0.29	-89.11
Santa Cruz	0.26	-87.01
San Diego	1.02	2.91
Santa Francisco	0.85	-27.36
USC	0.55	-72.01
US-CA	2.67	126.09
Non-US	0.44	-95.98
Other Institutions	1.72	44.83
Public Science	3.66	120.76
Grant year 88-90	1.02	0.95
Grant year 91-93	1.02	0.64
Grant year 94-96	1.36	10.32
Grant year 97-99	2.09	18.75
Paper pub year 85-88	1.30	26.76
Paper pub year 89-92	1.40	21.28
Paper pub year 93-97	1.10	4.78
β_1 **(obsolescence)**	0.12	75.20
β_2 **(diffusion)**	1.05E-08	10.06
Adjusted R-squared	0.220	
Number of observations	834624	

Figure V Fitted Citation Frequency (Base Category)



Source: Author's calculations.

Figure VI Average Science Citations Per Patent



Source: National Science and Engineering Indicators, 2002, National Science Foundation

Figure VII Citations to UC Berkeley Papers, US

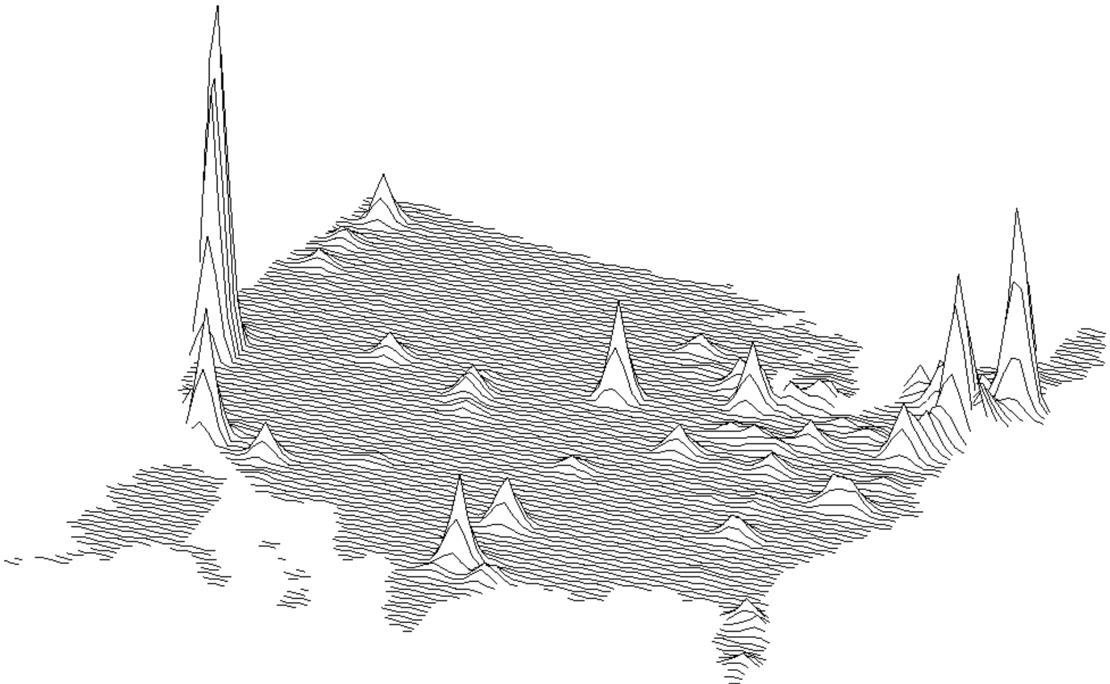


Table II Citation Function Results (excluding the patents applied in May/June 1995)

Variable	Coefficient	T-statistic for H_0: Parameter=1
Computers and Communications	0.04	-149.07
Drugs/medicine	2.29	86.30
Electronics	0.06	-157.78
Mechanical	0.01	-132.19
Other	0.05	-119.16
Biomedical research	40.67	9.43
Chemistry	5.59	7.83
Clinical Medicine	5.23	7.75
Eng/Technology	0.29	-6.99
Other Science	0.38	-6.37
Physics	0.61	-3.57
Caltech	1.15	18.36
Berkeley	0.58	-73.34
Davis	0.36	-118.03
Irvine	0.47	-82.48
Los Angeles	0.39	-119.34
Riverside	0.29	-98.94
Santa Barbara	0.30	-82.12
Santa Cruz	0.22	-85.19
San Diego	1.04	5.86
San Francisco	0.84	-26.02
USC	0.55	-67.73
US-CA	2.67	117.89
Non-US	0.45	-88.60
Other Institutions	1.72	37.49
Public Science	4.28	107.11
Grant year 88-90	1.03	1.32
Grant year 91-93	1.05	2.04
Grant year 94-96	1.42	11.54
Grant year 97-99	1.96	17.28
Paper pub year 85-88	1.31	25.15
Paper pub year 89-92	1.38	19.02
Paper pub year 93-97	1.16	6.85
β_1 **(obsolescence)**	0.122	73.44
β_2 **(diffusion)**	9.89E-09	9.42
Adjusted R-squared	0.195	
Number of observations	834624	

Base categories: Patent technology category=chemicals, scientific field=biology, academic institute=Stanford, patent assignee location=U.S./non-California, patent assignee type=firm, patent grant year=[1983, 1987], paper publication year=[1981, 1984]

Table III Citation Function Results
Full sample (excluding the patents granted to public science institutions, and the patents applied in May/June 1995)

	Coefficient	T-statistic for H_0: Parameter=1
Computers and Communications	0.039	-65.75
Drugs/medicine	4.400	44.20
Electronics	0.070	-58.35
Mechanical	0.024	-54.62
Other	0.017	-57.01
Biomedical research	82.019	2.89
Chemistry	2.559	1.69
Clinical Medicine	8.810	2.59
Engineering and Technology	0.396	-1.86
Other Science	0.502	-1.56
Physics	0.680	-0.90
Caltech	1.189	16.67
Berkeley	0.399	-84.70
Davis	0.206	-113.93
Irvine	0.242	-89.45
Los Angeles	0.333	-98.85
Santa Barbara	0.609	-32.62
Riverside	0.163	-88.64
Santa Cruz	0.181	-66.39
San Diego	0.533	-62.41
San Francisco	0.509	-72.20
USC	0.599	-43.47
US-CA	2.696	67.36
Non-US	0.375	-74.64
Other Institutions	2.196	74.20
Grant year 88-90	1.204	6.08
Grant year 91-93	1.044	1.38
Grant year 94-96	1.286	6.53
Grant year 97-99	2.221	14.03
Paper pub year 85-88	0.740	-25.99
Paper pub year 89-92	0.540	-36.42
Paper pub year 93-97	0.413	-40.76
β_1 **(obsolescence)**	0.123	45.24
β_2 **(diffusion)**	5.38E-09	2.92
Adjusted R-squared	0.118	
Number of observations	556416	

Base categories: Patent technology category=chemicals, scientific field=biology, academic institute=Stanford, patent assignee location=U.S./non-California, patent assignee type=firm, patent grant year=[1983, 1987], paper publication year=[1981, 1984]

Table IV Wald Tests of Restrictions

Hypotheses:
(1) H_0: All coefficients of patent technology categories are the same.
(2) H_0: All coefficients of paper fields are the same.
(3) H_0: All coefficients of patent grant years are the same.
(4) H_0: All coefficients of paper publication years are the same.

Hypothesis	*Test results*	
	Chi-Sq. (p-value)	*Adj. R^2 (rest.)*
(1)	36983.7 (0.000)	0.081
(2)	656.1 (0.000)	0.051
(1) and (2)	37639.9 (0.000)	0.028
(3)	1256.6 (0.000)	0.193
(4)	1435.8 (0.000)	0.193
Adj. R^2 (unrest).	0.195	
# of obs.	834624	

Table V
Bio Nexus Results

Variable	Coefficient	T-statistic for H_0: Parameter=1
Drugs & Medical	2.394	36.49
Chemistry	0.192	-68.28
Clinical Medicine	0.183	-96.74
Other Biotech	2.325	41.93
Caltech	0.900	-6.36
Berkeley	0.474	-45.54
Davis	0.303	-61.64
Irvine	0.367	-46.95
Los Angeles	0.364	-57.15
Riverside	0.222	-50.74
Santa Barbara	0.264	-39.76
Santa Cruz	0.241	-36.67
San Diego	0.980	-1.35
San Francisco	0.778	-17.64
USC	0.506	-33.72
US-CA	2.541	49.27
Non-US	0.443	-38.89
Other Institutions	1.586	13.50
Public Science	4.153	45.76
Grant year 88-90	1.027	0.48
Grant year 91-93	1.021	0.38
Grant year 94-96	1.364	4.57
Grant year 97-99	1.703	6.37
Paper pub year 85-88	1.298	10.68
Paper pub year 89-92	1.313	7.04
Paper pub year 93-97	1.102	1.91
β_1 **(obsolescence)**	0.113	28.78
β_2 **(diffusion)**	3.23E-07	18.17
Adjusted R-squared	0.184	
Number of observations	158976	

Base categories: Patent technology category=chemicals, scientific field=biochemistry, biophysics, and molecular biology, academic institute=Stanford, patent assignee location=U.S./non-California, patent assignee type=firm, patent grant year=[1983, 1987], paper publication year=[1981, 1984]

Table VI
Non-bio Results

Variable	Coefficient	T-statistic for H_0: Parameter=1
Computers & Communications	17.602	8.66
Electronics	7.185	7.86
Mechanical	4.539	6.94
Other	0.207	-5.50
Eng/Technology	9.289	3.07
Other Science	0.520	-1.54
Physics	44.397	3.38
Caltech	0.142	-133.39
Berkeley	0.089	-139.61
Davis	0.454	-53.52
Irvine	0.011	-105.88
Los Angeles	0.044	-131.93
Riverside	0.013	-85.92
Santa Barbara	0.039	-133.34
Santa Cruz	0.005	-81.37
San Diego	0.025	-121.39
San Francisco	0.002	-139.44
USC	0.066	-113.47
US-CA	21.629	9.85
Non-US	0.781	-1.66
Other Institutions	0.469	-1.81
Public Science	38.070	8.53
Grant year 88-90	16.349	20.87
Grant year 91-93	7.158	14.52
Grant year 94-96	15.722	13.37
Grant year 97-99	16.127	8.71
Paper pub year 85-88	0.210	-181.91
Paper pub year 89-92	0.081	-206.58
Paper pub year 93-97	0.022	-382.75
β_1 **(obsolescence)**	0.542	106.29
β_2 **(diffusion)**	5.81E-11	2.89
Adjusted R-squared	0.069	
Number of observations	397440	

Base categories: Patent technology category=chemicals, scientific field=chemistry, academic institute=Stanford, patent assignee location=U.S./non-California, patent assignee type=firm, patent grant year=[1983, 1987], paper publication year=[1981, 1984]