Cite Seeing: Patent Citations and the Economic Value of Patents

Bhaven N. Sampat
School of Public Policy
D.M. Smith Building
Georgia Institute of Technology
Atlanta, GA 30332 USA
e-mail: bhaven.sampat@pubpolicy.gatech.edu

Arvids A. Ziedonis
University of Michigan Business School
701 Tappan Street
Ann Arbor, MI 48109-1234 USA
e-mail: azied@umich.edu

This Draft: November 1, 2002

ABSTRACT: Over the last decade, the use of patent citation based measures as proxies for the private value of patented technologies (usually an unobserved variable) has become popular in the economics of innovation, management of technology, and applied industrial organization literatures. This paper assesses the degree to which citations are related to private value in a dataset where the latter is observed. In particular, we use data on patenting and licensing by two major research universities to assess whether citations can predict whether university technologies are licensed and the amount of revenues they earn if licensed. We find that citations are good indicators of whether a patent is licensed. However, they are not good indicators of revenues earned upon licensing, a result that is robust across different estimation procedures.

JEL Classifications: O31, O32
Keywords: Intellectual property, patents, licensing, university-industry linkages, knowledge transfer

The authors gratefully acknowledge support from Columbia University and the Mack Center for Technological Innovation at the Wharton School of the University of Pennsylvania. We thank Rebecca Henderson, David Mowery, Richard Nelson, and participants at the NBER productivity meeting for insightful comments on earlier versions of this research. Lynn Fissell generously provided data used in this study.

1. Introduction

With the growing recognition of the economic importance of technological innovation, demand by economists and other scholars for measures of inventive outputs has increased dramatically over the past two decades. Measurement of such outputs has been frustrated, however, because key theoretical constructs such as “technological advance” and “knowledge spillovers” are not directly observable and thus difficult to quantify.[1] Because they are widely available in electronic form however, patent data have increasingly been employed to construct proxy variables for these unobserved concepts (e.g., Jaffe 1998; Jaffe and Trajtenberg 2002). The literature has made heavy use of patent citations (i.e, citations by patents to previous patents as prior art) to estimate knowledge flows, and patent citation counts have been used to estimate both the private and the social value of patented technologies. Many of these studies have been unable to construct direct measurement of economic value, relying instead on estimates of social surplus (Trajtenberg 1990), calculations of Tobin’s Q (Hall, et al. 2000), or surveys of patent holders (Harhoff et al. 1999a, 1999b), for example.

In this paper, we investigate the relationship between revenues generated from the licensing of patents by two major research universities and the pattern of citations that these universities’ patents receive. Our goal is to assess the degree to which patent citation counts are “good” proxies for the private value of patented inventions. Validating citation based measures is difficult because of the dilemma posed by Trajtenberg et al. (1997), who ask: “how can we establish the connection between a candidate proxy and [unobserved variable] x^*, given that by

[1] Krugman (1991) has noted “knowledge flows … are invisible; they leave no paper trail by which they may be measured and tracked, and there is nothing to prevent the theorist from assuming anything about them that she likes” (p. 53). Measuring innovation is an important topic not only in the economics literature proper, but amongst scholars of the management of innovation. Richard Rosenbloom recently suggested “It’s the holy grail of people working on the management of technology - being able to measure innovation…I don't think anybody has cracked it, not yet” (quoted in Buderi 2000).

definition no direct data exist on x^*" (31)[2]. Because so little is known about the relationship between patent citations and *total* economic value, we strive for a more modest goal - to examine whether citations can predict *private* value for a sample of patents where data on the latter (i.e. x^*) are available.

This paper is organized as follows. Section 2 briefly discusses how scholars have interpreted patent citations and reviews previous validation studies of citations as economic indicators. In Section 3, we describe the dataset used in this study. Section 4 describes our econometric methodology and presents the main results. To tip our hand, we find that while citations are good predictors of *whether* a technology is licensed, they are not significantly related to the *level* of license revenues earned conditional on licensing. Section 5 concludes with a discussion of these results and implications for future research.

2. The Uses of Patent Citations in Economics

Patent citations and patent citation based data have been employed in studies of innovation in response to limitations with the use of "simple" aggregate patent counts as measures of innovative output (see Griliches 1990 for a review). The large variance in the economic and technological significance of individual patents renders simple patent counts as extremely noisy indicators of the innovative output of a firm or government program. A potential solution, suggested by Trajtenberg (1990), was to weight patents by the number of times they are cited in subsequent patents.

Since that work, economists have used citation counts to measure two different "values" of patented inventions. One is the patent's "social value," based on the assumption that a patent

[2] The authors point out that this question is not often asked in economics. See however Griliches (1983) and Klepper and Leamer (1984) on evaluation of proxy variables from an errors-in-variables framework, and Krasner and Pratt (1986) for a more general consideration.

A citing a patent B implies that A draws upon on the knowledge embodied in B, or that B is a technological antecedent of A (Trajtenberg et al 1997; Caballero and Jaffe 1993; Jaffe and Trajtenberg 1999). Citation of patent B by many subsequent patents suggests that numerous developments build upon the knowledge embodied in patent B (i.e., patent B has generated significant technological spillovers). Based on the rationale that inventions that generate a higher level of spillovers are more economically or technologically important, a significant stream of research has used citation counts to patents to assess social value (or "importance") of patented inventions (see Jaffe 1998 for a review).

In many of the seminal articles, the "building" metaphor is supported by analogy to bibliographic citations (i.e. citations in academic journal articles, which indicate the sources upon which an author relies).[3] Indeed, some suggest that patent citations are better indicators of sources upon which new knowledge builds than literature citations. Trajtenberg et al. (1997) argue that "[b]ecause of the role of the examiner and the legal significance of patent citations, there is reason to believe that patent citations are less likely to be contaminated by extraneous motives in the decision of what to cite than other bibliographic data" (XX).

The "building" metaphor may not be necessarily correct, however. Certainly, some of the previous patents upon which an invention builds (or technological antecedents, to use the language of Jaffe et al. 1998) will be "material to patentability" and cited, but not all will.[4] And there are certainly some items that a patent cites because they are "material to patentability," but upon which the invention did not build. These include not only citations inserted by examiners or the attorney that the inventor(s) did not know about when developing the invention, but also

[3] This is of course a gross oversimplification, and there has been considerable debate in the sociology of science and bibliometrics as to the "meaning" of bibliographic citations. See Melkers (1993) and Cole (2000) for recent overviews.
[4] Of course, many (and arguably most) of the sources upon which inventions build are not patented.

very similar inventions which the inventor knew about but did not build upon or draw on a knowledge spillover from, in the economic sense described by Griliches (1992).

Thus, the theoretical foundation for using citations as proxies for spillovers and social value may be weaker than conventionally thought, suggesting the need for empirical validation studies. Several recent papers by Adam Jaffe and colleagues (Jaffe, Trajtenberg, and Fogarty 2000; Jaffe, Fogarty, and Banks 1998) have found some support for the use of citation data for these purposes, but clearly more work remains to be done.

In addition to using citations as proxies for spillovers and citation counts as proxies for social value, scholars have also used citation counts as proxies for measures of the private value of the invention to the patent holder. This is essentially the maintained hypothesis of the various studies that have attempted to account for the "value" of a firm based on its citation-weighted patent stock (Hall et al. 2000; Shane and Klock 1997; Austin 1994). Citations have also been used as measures of the value of inventions in studies of why patents are litigated (Lanjouw and Schankerman 2000).

Interestingly, the literature is virtually silent regarding causal pathways through which citation counts might be related to the private value of inventions. Rather, the implicit view is that social and private values of patented inventions are somehow correlated, which although likely true, begs for further exploration.[5] We propose four possible causal channels through which citation counts could be related to the private value of inventions:

---

[5] Henderson et al. (1998) implicitly suggest such a correlation. While their discussion of the diminished "quality" of university patents after 1980 appears to reflect a social value construct (i.e., spillovers or the degree to which other inventions build upon university patents), the authors also suggest that a possible reason for this finding is that it was less costly for universities to patent after 1980, which effectively lowered the threshold quality level of invention above which they would file for patent protection. If universities are comparing private costs with private benefits, as they appear to be in this explanation, then increased patenting would only occur if the "quality" of the invention was correlated with the expected private value of the patent.

Theory I: *Citations Reflect the Portion of Social Returns Appropriated*. Under this theory, citations to a patented invention represent knowledge flows or spillovers, and citation counts reflect the social value of inventions, as discussed above. However, patent holders are able to appropriate a sufficient portion of the social returns so that citations are also useful as a proxy for the private value of inventions.

Theory II: *Citations Reflect Entry into Profitable Areas of Research*. A patent that has been revealed to be profitable will induce other firms to undertake research in technologically close but non-infringing areas, (probabilistically) resulting in citing patents.[6]

Theory III: *Citations Indicate the Technological Opportunities or Market Interest in a Technological Area*. Citations reflect the presence of commercial interest and activity in a field, and hence profit potential.

Theory IV: *Citations Result from a Public Disclosure*. Patents that are economically successful are more widely known and more likely to be cited as prior art. Lanjouw and Schankerman (2000) term this a "publicity effect."[7]

In principle, detailed data on the timing of citations and identity of citers could help to distinguish which of these causal pathways (if any) are responsible for the citations-private value relationship. Theory I has implications for the identity of citers: it suggests that a significant share of spillovers are appropriated, and that these appropriated spillovers account for private value to the patent holder. Under this theory, a significant share of citations may be generated by parties that are likely to compensate the patent holder (e.g. licensees), and such citations should

[6] This is also suggested by Trajtenberg (1990): "The very existence of those later patents attests to the fact that the earlier patents opened the way to a technologically successful line of innovation. More important, they presumably attest to the *economic* success as well (at least in expected value terms), for those patents are the results of subsequent costly innovational efforts undertaken mostly by profit seeking agents ... If citations keep coming, it must be that the innovation originating in the citing patent had indeed proved to be valuable" (189).

[7] The authors find that citations increase immediately after litigation. They suggest that this is because once patents are litigated, they are more well known to examiners and applicants, and thus more likely to be cited.

be more closely related to economic value than citations by others. Theories II and IV have implications about the timing of citations relative to realizations of private economic value, (e.g. licensing or profitability). In particular, they suggest that realizations of private returns will precede citations. Thus, under theories II and IV, very early citations would not be correlated with private economic value, while later citation counts would be. Below, we attempt to distinguish between these hypotheses using data on the timing and identity of citing relative to licensing.

The main objective of our study, however, is to assess whether citations are related to measures of private value. This exercise is similar in spirit to several recent attempts to validate citation counts as measures of private economic value. Some authors have used measures of value at the firm level, considering whether a firm's citation-weighted patent stock appears to impact its market value (Deng et al. 1999, Shane and Klock 1997, Austin 1993, Hall et al. 2000). The most comprehensive of these studies is Hall et al. (2000), which finds a significant relationship between the "Tobin's Q" of firms and their citation weighted knowledge stocks and other inputs. Other scholars have focused on the individual patent as the unit of observation. Lanjouw and Schankerman (1999), using a latent variable model, and find that citations are positively correlated with other measures of the value of patents.

In work more closely related to this study, Harhoff et al. (1999a, 1999b) find that citation counts appear to reflect the "asset value" of patents, or the price at which surveyed patent owners reported they would be willing to selling the rights to particular patents. In the only other study of which we are aware in which the authors had access to direct measures of private returns from patents, Mogee et al. (1997) compare citation counts to various measures of the value of a sample of patents, including an estimate of patent value from a patent renewal model, whether a patent is licensed, and the amount of license revenues earned by a patent. They find that the number citations are positively and significantly related to renewal model value estimates and whether a patent is licensed, but that there is no significant relationship between citations and the level of revenues.

3. Data

We utilize data on *university* patents, citations, license outcomes, and license revenues to examine the citations-value relationship. After the passage of the Bayh-Dole Act of 1980, universities became more active in patenting and licensing inventions generated by faculty research (Henderson et al. 1998; Mowery et al. 2001). One advantage of using university data to study the relationship between patent citations and economic value is that unlike the private sector, the university lacks the requisite complementary assets and the motive to engage in product development and marketing activities to capture economic value. Therefore, universities typically apply for patent protection solely for the purposes of licensing inventions generated by research.[8] Absent the strategic motives to patent to prevent competitors from using the new technologies or to block competitors from patenting, the private value to a university of a patent right (or the "asset value") derives primarily from licensing revenues.[9] Another advantage of utilizing university data is that few private firms are able to patent in as wide a range of fields as

[8] However, Mowery and Sampat (2001b) note that earlier in the 20th century, universities had other motivations for patenting as well, including preventing firms from patenting the fruits of university research and monopolizing an emerging technological field; and to assure that only "reputable" producers exploited university inventions, thus protecting universities from "bad press." As a result, university patents were often dedicated to the public. However, since the 1970s and especially since the passage of the Bayh-Dole act of 1980, a primary motive for university patenting has been to license inventions.

[9] Invention value to the University may also be derived not only from its licensing revenue, but also from fostering closer interaction with licensees, which may yield future benefits in the form of research funding or jobs for students. To this extent, *whether* an invention is licensed may be more appropriately measure the "private value" of an invention to the university than license revenues *per se*.

are major research universities. University data therefore enable us to examine the relationship between citations and revenues across several broad technological fields.

On the other hand, universities' licensing strategies may have objectives other than maximizing the license revenues from patents. In particular, the Bayh-Dole Act called on universities to locate licensees that are best able to commercialize inventions, which may not necessarily be those that could generate the highest royalty income. Moreover, university patents may themselves be qualitatively different from corporate patents (Jaffe et al. 1993, Trajtenberg et al. 1997), which may limit the degree to which out results are generalizable. We revisited these issues below.

Our sample of patents was generated from archival data at the technology transfer offices of Columbia University and the University of California. These data contain disclosures of inventions made by faculty, researchers, students, and staff at these two universities, and the patenting and licensing outcomes of these invention disclosures. Our sample contains 1208 issued patents applied for by the University of California and Columbia University between 1980 and 1994. We observe whether each of these patents was licensed by the end of 1999; if licensed, the identity of the licensee, and the total amount of dollar payments made to the university by that date. We also utilize the Micropatent database of US patents to identify all subsequent patents that cite our sample patents as prior art by the end of 1999. Table 1 shows trends in total patents issued to the two universities by application year. The University of California (UC) accounts for 85% of the patents in our sample, reflecting its greater number of campuses (nine vs. two) and its long history of involvement in patenting - Columbia gradually entered into patenting activities only after the passage of the Bayh-Dole Act, in 1980 (Mowery et al. 2001; Mowery, Sampat, and Ziedonis 2001).

*** Insert Table 1 Here ***

The dependent variables in our analysis are licensing outcomes, (a) whether a patent is licensed, and (b) the revenues it earns conditional on licensing. Table 2 shows the distribution over time of the first of these variables (whether licensed) for the pooled sample of patents. This table reports the proportion of patents in each application year that were licensed by 1999. Within application years and in the overall sample, fewer than half of the patents are ever licensed. Since licensing is the one of the only means through which universities reap returns from patents,[10] this table suggests that less than half of university patents have the potential to earn revenues. This implies further that the value of university patents will have a skewed distributed - more on this later.

*** Insert Table 2 Here ***

One empirical problem we face is that the series is right censored (i.e., we only know if a patent has been licensed by a given point in time, not whether it will be ultimately licensed). For example, it may be that more than 41% of the patents in our sample applied for in 1994 will eventually be licensed, though they were not by 1999. We consider this characteristic of the data in developing the econometric setup, and in interpretation of results, below.

[10] Mowery and Sampat (2001a) show that in the early half of the century, important motives for university patenting included prevention of "patent piracy" (patenting of university inventions by others) and "quality control" of users of university research outputs.

More so than whether a patent is licensed, the licensing revenues technologies earn may indicate the value of patented inventions to universities.[11] Several issues complicate measurement of patent licensing revenue. First, patents are often licensed in bundles ("inventions"), and revenues accrue to the inventions or groups of inventions rather than to individual patents. Because we are not able to observe the relative importance of each patent in a bundle, we have no rule to allocate licensing revenues across bundled patents. We therefore use licensed inventions as the unit of analysis in our licensing income specification. A second issue concerns the types of revenues to include as "license revenues." Licensees may pay up-front fees upon execution of the contract, annual fees to keep the license active, milestone payments based on level of sales or other events (e.g., reaching some stage of clinical trials), sales based royalties, as well as legal reimbursements and other fees. Since we are interested in the private value of the patented invention to the university (similar to the asset value of the patent to a firm lacking strategic motives, see Harhoff et al. 1999), we use all revenues except for reimbursements.[12] A third issue is whether to treat unlicensed patents as observations with zero revenues or to exclude them from the analysis relating licensing income to citations. We chose the latter strategy, both because the "technology" unit of analysis is not properly defined for unlicensed inventions (patents are bundled at licensing) and because it allows for distinguishing between licensed inventions with zero revenues and unlicensed inventions.

Like the data on whether a patent is licensed, revenue data are right-censored -they are subject to truncation bias; at any given time we observe only a fraction of the lifetime revenues earned by licensed inventions. The resulting data on gross revenues for licensed technologies are

[11] Jensen and Thursby's (1998) report on a survey of major university technology transfer offices indicates that the most important goal for university technology transfer officers is license income.
[12] Additional analyses (unreported) show that the main qualitative results are not affected if we look at sales-based royalties alone.

also extremely skewed: in 1980 the top 10% of licensed technologies account for 95% of total gross income, and in 1990 the top 10% account for 88%. This feature of the distribution - that outlying tail values account for a large proportion of cumulative revenue - is consistent with previous evidence on the distribution of returns from industrial innovations (see Scherer and Harhoff 2000 for an excellent review), and university inventions (Mowery et al. 2001; Mowery and Sampat 2001). As is common practice in studies employing highly skewed data therefore, we use a log-transformation of gross revenues as the dependent variable. Because many of the licensed technologies within the sample earn no revenues, we construct the dependent variable as the log of $1 plus license revenues. The charts in Figure 1 show the distribution of this variable in each application year.

*** Insert Figure 1 Here ***

Turning attention now to the main independent variable, patent citations, Figure 2 shows the distribution of citations by application year at the level of the patent. The patent citation distribution is also extremely skewed, as would be expected if patent citations were related to the value of inventions. (The distribution of unique citations at the level of technologies is similarly skewed.)

*** Insert Figure 2 Here ***

As mentioned earlier, one goal of this paper is to use these data to shed light on the causal channels through which citations and value may be related. Table 3 reports the citations made by licensees as a fraction of overall citations by application year of our sample patents. Overall, licensees account for a small share of citations (16%), though there is considerable variation over time.[13] Since only licensee citations are associated with compensation to the university, this low number suggests that Theory I above - that citations reflect social value but that part of this is appropriated by patent holders - is unlikely to be the main or only causal link between citations and value, at least in this sample. To consider this issue more carefully, we examine the relationships between licensee citations and revenues from licensed inventions below.

The third column of Table 3 shows that overall, a low proportion of citations are made before the license execution date. This provides some support for Theories II and IV, under which citations follow economic success. We revisit this suggestion below.

*** Insert Table 3 Here ***

Before turning to the econometric analysis, it is useful to look at the basic relationships between the indicators of economic value and citations. Table 4 reports the mean number of citations for licensed and unlicensed patents. Both means decline over time, likely an artifact of truncation bias. In each application year however, licensed patents have a higher number of citations, on average, though the magnitude of the difference varies over time. Indeed, since both patent citation and the licensing series are truncated, we can ask whether *at a given point in time* citations are informative of the license status of an invention or the revenues earned by an invention. Figure 3 shows a scatter plot of log revenues versus citations for all of the licensed

[13] Firms (especially multi-divisional firms) often hold patents in the names of their subsidiaries or their parent firms (see Hall et al. 2000). To ensure that we captured all citations made by licensees or their corporate parents or subsidiaries, we utilized the *Directory of Corporate Affiliations* to identify the corporate affiliations of the assignee

technologies. These data provide preliminary evidence that there is a positive relationship between revenues and citations, though the clusters of points along the *x* and *y* axes (uncited patents with revenues, cited patents without revenues) also suggest that the relationship is noisy.

*** Insert Table 4 and Figure 3 Here ***

In the next section, we explore the relationship between citations, licensing, and license revenues more systematically, controlling for technology field and time effects.

4. Econometric Methodology and Results

To examine whether citations can predict whether a technology is licensed, we estimated probit models regressing a dummy variable indicating if the patent was licensed on citations, technological field controls, and application year controls. Results are reported in Table 5. The positive and statistically significant coefficient for total citations (Model 1) indicates that citations have are related to whether a technology is licensed. Probit coefficients give the effect of a one unit change in X on the cumulative normal probability of Y, and are thus difficult to interpret directly. We therefore calculate the marginal effect of citations on the probability of licensing (calculated at the mean of the data). The marginal effect is approximately .007 (i.e., an additional citation leads to a .7% increase in the probability that a patent is licensed). Figure 4 plots the predicted marginal effects across the range of the citations variable: note that the marginal effects are greatest (about 0.74%) at about 20 citations.

*** Insert Figure 4 Here ***

---

of each citing patent. If the licensee was a subsidiary or parent of the citing firm, we considered that citation as

Because it is not strictly correct to speak of infinitesimal changes in integers such as citation counts (Caudill and Jackson 1989), we also we calculate the predicted probabilities of a patent being licensed as a function of the number of citations. At the mean level of citations (8 citations), the predicted probability of licensing is 41.2%, and at the median (4 citations) the predicted probability is 35.4%. Increasing citations from the median level to the 75th percentile level (9 citations) increases the probability of licensing by 3.6%, and increasing citations from the median to the 95th percentile level (28 citations) increases the probability by 17.6%. Figure 5 shows the predicted probabilities of licensing by the number of citations a technology receives.

*** Insert Figure 5 Here ***

Figure 5 shows that as we approach the right tail of the citations distribution, the probability of licensing approaches 100%. The slope of the predicted probabilities is greatest for low values of citations (see also the marginal effects in Figure 4), suggesting that there are diminishing informational value to citations. To test this more formally, in separate regressions we include citations-squared as an independent variable. The quadratic term waiss negative and statistically significant (at the 1% level), (consistent with diminishing informational content) but extremely small. The implied full marginal effect of citations does not change dramatically, except for extremely large values of citations. Consequently, we exclud the quadratic term from the remaining specifications.

To check for possible differences in the citations-licensing relationship across universities, we estimate the baseline model for the University of California patents alone and

---

being made by the licensee.

the Columbia patents alone. The results of these separate regressions are reported in Models 2 and 3 of Table 5. The probit coefficients for total citations are .017 and .029 respectively. Model 4 includes a term interacting citations with the Columbia dummy: the coefficient on this term thus gives differences across the universities in the effect of citations. The coefficient on the interaction term is small and statistically insignificant: we cannot therefore reject the hypothesis that the citations-licensing relationship is the same across the universities.

The probit results in Models 1-3 thus suggest a statistically and qualitatively significant relationship between citations and whether a patent is licensed. Citations provide information on whether an invention is licensed above and beyond what we can infer from time effects and technology class effects alone. These results are consistent with licensing policies practiced at Columbia and UC. Our interviews with licensing officers suggest that at these two universities license all patents when it is possible to find a licensee. Moreover, these universities frequently require firms to underwrite patent application costs in exchange for a license. Thus, the evidence thus far suggests that citations are good indicators of whether there has been some commercial interest in a patent, consistent with Theory III above.

Of course, licensed patents are only "valuable" to a university in an expected value sense. In order to examine whether citations reflect (or can predict) the actual returns from licensed inventions, we estimate tobit regressions analogous to the probit regressions above. Recall, as explained above, that in our analysis of licensing revenues, our unit of analysis is a licensed invention disclosure, or technology, rather than a patent. Tobit regressions are appropriate in this context because the dependent variable, (log of $1 plus license revenues), is bounded below by zero. To be more precise, the source of the censoring is that among licensed technologies, for all observations for which a latent variable (say "quality") is not sufficiently high, we will observe zero revenues.

The results from the tobit regressions are presented in Table 6: Model 5 gives results for the entire sample, Model 6 for the University of California only, Model 7 for Columbia, and Model 8 includes the interaction term. For the overall sample, the coefficients imply a marginal effect of 0.022, i.e. a one unit change in citations leads to a 2.2% increase in revenues.[14] However, though qualitatively significant, the effect is not statistically significant from zero in any of the models. These results suggest that citations do not appear to be good predictors of revenues earned by licensed technologies.

** Insert Table 6 Here ***

Though the tobit results provide little evidence of a relationship between citations and license revenues, these results may reflect the sensitivity of tobit specifications to violation of various statistical assumptions (Maddala 1983). Of particular concern are heteroskedasticity[15] and the normality of residuals; the tobit model is completely identified by these assumptions. In absence of these conditions, tobit estimates are inconsistent and inefficient (Maddala 1983; Deaton 1997; Arabmazar and Schmidt 1982).

We therefore also specify a non-parametric regression analogous to the tobit model, but robust to violations of heteroskedasticity and non-normality assumptions: Powell's (1984) censored least absolute deviations (CLAD) estimator. The CLAD is a special case of the least

[14] Marginal effects are calculated at the mean profile, using the procedure suggested by McDonald and Moffitt (1980).

[15] It is likely that there is more room for idiosyncracies in license revenues for low quality (and under the maintained hypothesis, heavily cited) inventions.

absolute deviation estimator (LAD), or the median regression estimator; median regression is in turn a special case of quantile regression. Quantile regression, pioneered by Koenker and Bassett (1978), estimates conditional quantile functions rather than conditional mean functions. It is typically employed when researchers are interested in the effects of covariates on the entire distribution of the dependent variable, rather than the mean alone. Coefficients from median regressions estimate the effect of independent variables on the conditional median of the dependent variable.

Estimation of the conditional median function amounts to minimizing the sum of absolute residuals, or least absolute deviations(LAD) (Deaton 1997). Powell (1984) developed a method for estimating median regressions in the context of non-negativity constraints and censored data: CLAD regressions. Moreover, CLAD estimators are "distribution free" in the sense that no strong assumptions are needed on the distribution of the residuals, making them particular suitable for contexts where we might suspect heteroskedasticity or non-normality of residuals. In the present context, the CLAD estimator has the additional benefit that it is less sensitive to extreme outliers than the tobit estimator (since the conditional median is less sensitive to outliers than the conditional mean).

To generate CLAD regression estimates, we use an algorithm suggested by Buchinsky (1994). The essence of this procedure is to estimate the LAD regression for all observations, discard all observations where predicted values are negative, and repeat the last step until all estimates converge.[16] (Convergence in our sample was typically achieved within 25 iterations.) Following the suggestion of Deaton, we report bootstrapped standard errors from the final

[16] The algorithm was programmed into Stata using an adaptation of code developed by Angus Deaton, and discussed in Deaton (1997). It is available from the authors upon request.

iteration, since conventional standard errors for median regression estimates (Koenker and Bassett 1984) are unreliable in the context of heteroskedasticity.

The CLAD regression results are presented in Table 7. Interestingly, the coefficients from these regressions are similar to those obtained from the tobit regressions earlier. Moreover, they remain statistically insignificant at conventional levels, in the overall sample and in each of the sub-samples. Accordingly, it appears that the basic results of the tobit estimations - that citations are not good predictors of revenues for licensed technologies - are thus confirmed by this more robust estimation technique.

*** Insert Table 7 Here ***

Though citations are not related to license revenues in this sample, it is possible that citations by different types of citing entities are related to license revenues (along the lines suggested by the “theories of citation” posited above) and that their effect is being muddled by the other citations. To examine this possibility, we disaggregated our sample by (a) pre-license citations versus post-license citations, and (b) by citations from licensees and citations from others. We estimate tobit and CLAD regressions where we allowed these types of citations to enter separately. Results from the tobit specification are presented in Table 8 (Models 14 and 15), and results from the CLAD specification are presented in Table 9 (Models 16 and 17). Note that the coefficients are not stable across the models, making it difficult to interpret the qualitative effects of these different types of citations. More importantly, as with the pooled models, none of these “types” of citations have a statistically significant impact on revenues.

*** Insert Tables 8 and 9 Here ***

The above results suggest that citations are not related to license revenues, for licensed technologies. It is possible, however, that these results reflect truncation bias, if the citations-revenues relationship takes time to develop. The regressions above account for time effects in the sense that the coefficient on the citation variable gives the effect of citations on revenues, independent of the effect of time on revenues. A more subtle possibility is that there is a citations-revenue relationship, but it is only observable after a sufficient period of time. If so, inclusion of later observations for which the relationship has not had sufficient time to develop might obscure any citations-revenue relationship in earlier observations. To check for this, we estimated tobit and CLAD regressions where, in addition to the application year dummies and the technology class dummies, we included year dummies interacted with the citations term. In this specification, the coefficient on the interaction term would give the impact of citations within each application year. Table 10 shows the results of these estimations.

*** Insert Table 10 Here ***

The coefficient on the interaction term is insignificant for most application years, again suggesting that citations are not good predictors of revenues for licensed technologies, and that the results above are not driven by truncation bias.

5. Discussion and Conclusions

In the economics of innovation, many important concepts are typically unobservable or difficult to measure systematically. Patent citation counts have been used as proxies for one such concept, the private value of patents. In this paper, we asked whether citations could predict two direct measures of value, whether a patent is licensed and how much revenue it generates conditional on licensing, for a sample of university patents where these measures were observable.

The main finding of the paper is that while citations are good predictors of whether a university patent is licensed, they are not good predictors of the license revenues earned by technologies conditional on licensing. These results are broadly consistent with Mogee et al. (1997).

How should we interpret the result that citations are good predictors of licensing, but not of revenues conditional on licensing? One possibility, suggested by Theory III above, is that citations reflect market interest in areas in technological proximity to particular patents. Market interest induces innovative effort in particular technological areas, increasing the probability of later citations. At the same time, market interest also increases the probability of licensing. However, as innovation and commercialization are uncertain activities, the level of revenues ultimately earned by particular technologies may be influenced by factors other than market interest, including competition by competing technologies, licensees' commercialization incentives, and R&D and marketing competencies.

Thus we interpret the overall results as giving a preliminary nod to Theory III. The more fine-grained tests of the "theories of citation" reveal little evidence in favor of Theory I (citations reflect appropriability of spillovers), since licensees account for a small share of all citations, and it is difficult to imagine means other than licensing through which universities could appropriate social returns. The occurrence of most citations after the license is executed provides some support for Theories II (citations reflect entry into profitable areas) and IV (citations reflect a disclosure effect), the fact that these later citations are not related to commercial success suggest that these do not explain the entire story. Clearly, more work remains to be done on this front.

In assessing whether citations are "good" proxies, we have focused only on the qualitative and statistical significance of the coefficients in regressions of value measures on citations. Of course, there are other dimensions across which we might want to evaluate citations as proxies (see Krasker and Pratt 1986). From an errors-in-variables perspective, we would want to assure that the noise in citations (measurement error) is not related to the level of citations or the level of the underlying variable of interest. Such an analysis remains for future work.

We conclude by emphasizing the limits to the representativeness of this sample. University patents are very different from other patents (Trajtenberg et al. 1997) and universities motivations for patenting also vary (Mowery et al. 2001). Citations may be more or less closely related to the "value" of university patents than other patents - this is an open empirical question. If the reader prefers, she can consider this exercise a validation study of the relationship between citations and the value of university patents. The results here are material even under this more limited interpretation however, since an important use of citations in the economics literature has been as measures of the "quality" of university patents (see Jaffe and Trajtenberg 2002).

A decade ago, Griliches (1990) noted that the use of patent citation measures "is only in its beginnings and we are likely to see a much wider use of it in the future" (1689). As this statement remains true today, so does the need for more validation studies of the use of patent citation based measures

Tables and Figures

Table 1: Total Patents Assigned to the University of California and Columbia University, by Application Year

Application Year	University of California	Columbia University	Total
1980	55	2	57
1981	45	1	46
1982	50	5	55
1983	54	4	58
1984	57	9	66
1985	62	12	74
1986	60	12	72
1987	70	16	86
1988	76	21	97
1989	82	15	97
1990	72	12	84
1991	72	15	87
1992	76	17	93
1993	108	13	121
1994	91	24	115
Total	**1030**	**178**	**1208**

Table 2: Distribution of Licensing Outcomes by Application Year

Application Year	Proportion of Patents Unlicensed (%)	Proportion of Patents Licensed (%)
1980	73.68	26.32
1981	76.09	23.91
1982	60.00	40.00
1983	58.62	41.38
1984	62.12	37.88
1985	54.05	45.95
1986	52.78	47.22
1987	50.00	50.00
1988	51.55	48.45
1989	58.76	41.24
1990	58.33	41.67
1991	56.32	43.68
1992	60.22	39.78
1993	61.16	38.84
1994	59.13	40.87
Total	**58.69**	**41.31**

Table 3: Licensee Citations, Pre-License Citations for Licensed Patents by Application Year of Cited Patent

Application Year	Proportion of Citations By Licensees (%)	Proportion Citations Occurring Before Licensing (%)
1980	47	11
1982	13	15
1983	29	5
1984	9	2
1985	11	10
1986	7	18
1987	29	16
1988	3	40
1989	0	32
1990	14	19
1991	16	8
1992	15	12
1993	30	0
1994	15	0
1995	0	0
Overall	**16**	**14**

Table 4: Mean No. of Citations for Unlicensed, Licensed Patents

Application Year	Mean No. of Citations for Unlicensed Patents	Mean No. of Citations for Licensed Patents
1980	9.67	26.07
1981	7.17	15.55
1982	8.91	15.18
1983	12.44	16.42
1984	7.68	20.16
1985	10.40	13.74
1986	8.58	13.88
1987	7.91	8.84
1988	8.38	12.09
1989	5.72	7.00
1990	4.37	9.26
1991	3.96	6.82
1992	3.57	3.84
1993	1.66	2.11
1994	1.81	2.32
Overall	**6.16**	**9.81**

Table 5: Probit Estimation for UC and Columbia

	Dependent Variable: "Was the Patent Licensed?" (1=Yes, 0=No)			
	(1)	(2)	(3)	(4)
	Entire Sample	University of California Patents Only	Columbia Patents Only	Entire Sample, Include Interaction Term
Total Citations	0.019	0.017	0.030	0.019
	(5.36)**	(4.20)**	(3.54)**	(4.63)**
Citations * Columbia Dummy				0.000
				(0.05)
1981 Dummy	0.014	0.021		0.014
	(0.05)	(0.08)		(0.05)
1982 Dummy	0.356	0.387	0.951	0.357
	(1.39)	(1.47)	(0.47)	(1.39)
1983 Dummy	0.418	0.298		0.418
	(1.65)	(1.15)		(1.65)
1984 Dummy	0.303	0.202	2.323	0.304
	(1.23)	(0.79)	(1.13)	(1.23)
1985 Dummy	0.525	0.548	1.688	0.525
	(2.20)*	(2.21)*	(0.83)	(2.20)*
1986 Dummy	0.575	0.682	1.450	0.576
	(2.39)*	(2.71)**	(0.71)	(2.39)*
1987 Dummy	0.705	0.746	1.849	0.706
	(3.03)**	(3.07)**	(0.90)	(3.03)**
1988 Dummy	0.670	0.573	2.530	0.670
	(2.93)**	(2.39)*	(1.24)	(2.92)**
1989 Dummy	0.531	0.625	1.212	0.531
	(2.32)*	(2.64)**	(0.59)	(2.31)*
1990 Dummy	0.516	0.522	1.965	0.517
	(2.20)*	(2.14)*	(0.96)	(2.19)*
1991 Dummy	0.593	0.558	2.238	0.594
	(2.54)*	(2.29)*	(1.09)	(2.54)*
1992 Dummy	0.531	0.492	2.204	0.532
	(2.28)*	(2.02)*	(1.07)	(2.28)*
1993 Dummy	0.520	0.528	1.918	0.521
	(2.32)*	(2.29)*	(0.93)	(2.31)*
1994 Dummy	0.553	0.614	1.865	0.554
	(2.45)*	(2.60)**	(0.91)	(2.44)*
Chemicals Dummy	-0.162	-0.188	-0.034	-0.162
	(0.68)	(0.76)	(0.04)	(0.68)
Drugs and Medical Dummy	0.075	0.064	0.013	0.075
	(0.33)	(0.27)	(0.01)	(0.33)
Electronics Dummy	-0.080	0.009	-0.809	-0.080
	(0.34)	(0.04)	(0.84)	(0.33)
Mechanical Dummy	-0.676	-0.637	-1.076	-0.676
	(2.51)*	(2.26)*	(1.02)	(2.51)*
Constant	-0.796	-0.774	-2.295	-0.798
	(2.70)**	(2.55)*	(1.02)	(2.70)**
Observations	1205	1027	173	1205

Absolute value of *z*-statistics in parentheses
* significant at 5% level; ** significant at 1% level

Table 6: Tobit Estimation for Licensed Technologies Only

	Dependent Variable: Log(1 + License Revenues)			
	Entire Sample	University of California Technologies Only	Columbia Technologies Only	Entire Sample, Include Interaction Term
Total Citations	0.034	0.018	0.024	0.014
	(1.58)	(0.58)	(0.55)	(0.47)
Cites*Columbia Dummy				0.031
				(0.88)
1981 Dummy	-3.414	-3.594		-3.301
	(1.27)	(1.40)		(1.23)
1982 Dummy	-3.602	-2.881	-9.178	-3.553
	(1.60)	(1.28)	(1.22)	(1.58)
1983 Dummy	-6.180	-6.844	-5.810	-6.059
	(2.65)**	(2.93)**	(0.68)	(2.60)**
1984 Dummy	-1.287	-1.896	-2.759	-1.271
	(0.61)	(0.87)	(0.37)	(0.60)
1985 Dummy	-4.371	-3.798	-8.685	-4.380
	(2.09)*	(1.81)	(1.13)	(2.10)*
1986 Dummy	-7.142	-6.342	-14.006	-7.014
	(3.25)**	(2.95)**	(1.61)	(3.19)**
1987 Dummy	-3.770	-4.220	-3.099	-3.727
	(1.84)	(2.10)*	(0.37)	(1.82)
1988 Dummy	-2.630	-2.023	-8.053	-2.628
	(1.26)	(0.98)	(0.99)	(1.26)
1989 Dummy	-3.608	-3.640	-6.367	-3.598
	(1.71)	(1.78)	(0.73)	(1.71)
1990 Dummy	-4.233	-4.335	-7.974	-4.318
	(1.99)*	(2.04)*	(1.04)	(2.03)*
1991 Dummy	-6.754	-8.813	-3.453	-6.845
	(3.13)**	(3.97)**	(0.40)	(3.17)**
1992 Dummy	-9.067	-12.054	-5.735	-9.186
	(3.97)**	(4.94)**	(0.69)	(4.02)**
1993 Dummy	-9.104	-10.322	-5.288	-9.221
	(4.20)**	(4.80)**	(0.63)	(4.25)**
1994 Dummy	-12.725	-15.282	-3.199	-12.863
	(5.54)**	(6.43)**	(0.39)	(5.59)**
Chemicals Dummy	0.771	0.524	-0.993	0.651
	(0.34)	(0.23)	(0.16)	(0.29)
Drugs and Medical Dummy	-0.481	-1.054	-2.592	-0.605
	(0.22)	(0.47)	(0.44)	(0.28)
Electronics Dummy	-12.486	-13.245	-10.546	-12.574
	(5.32)**	(5.52)**	(1.80)	(5.36)**
Mechanical Dummy	-1.304	-1.939	-1.962	-1.433
	(0.49)	(0.72)	(0.24)	(0.54)
Constant	13.691	14.447	18.618	13.952
	(4.86)**	(5.12)**	(1.98)	(4.93)**
Observations	357	301	56	357

Absolute value of *t*-statistics in parentheses
* significant at 5% level; ** significant at 1% level

Table 7: CLAD Estimation for Licensed Technologies Only

	Dependent Variable: Log(1 + License Revenues)			
	Entire Sample	University of California Technologies Only	Columbia Technologies Only	Entire Sample, Include Interaction Term
Total Citations	0.038	0.018	0.032	0.015
	(1.86)	(1.09)	(0.53)	(1.29)
Cites*Columbia Dummy				0.030
				(1.52)
1981 Dummy	-1.304	-1.110		-1.588
	(0.73)	(0.42)		(0.87)
1982 Dummy	-2.136	-1.302	-1.156	-1.777
	(1.40)	(0.75)	(0.12)	(1.06)
1983 Dummy	-1.815	-1.608	-4.291	-1.748
	(1.22)	(0.75)	(0.55)	(0.94)
1984 Dummy	0.188	-0.445	-1.665	-0.386
	(0.12)	(0.27)	(0.30)	(0.26)
1985 Dummy	-0.497	0.075	-5.567	-0.338
	(0.45)	(0.04)	(0.71)	(0.19)
1986 Dummy	-1.831	-0.087	-6.454	-0.748
	(1.17)	(0.04)	(0.81)	(0.27)
1987 Dummy	-0.712	-0.302	-0.716	-0.624
	(0.54)	(0.16)	(0.09)	(0.35)
1988 Dummy	-0.636	-0.051	-4.625	-0.884
	(0.52)	(0.03)	(0.55)	(0.56)
1989 Dummy	-0.599	-0.015	-6.155	-0.509
	(0.52)	(0.01)	(1.08)	(0.25)
1990 Dummy	-0.670	0.039	-4.688	-1.021
	(0.43)	(0.02)	(0.50)	(0.56)
1991 Dummy	-0.905	-1.254	-3.458	-1.259
	(0.38)	(0.22)	(0.35)	(0.63)
1992 Dummy	-0.785	-11.049	-3.947	-11.456
	(0.16)	(2.25)*	(0.49)	(2.16)*
1993 Dummy	-2.560	-10.714	-3.262	-2.788
	(0.55)	(2.12)*	(0.50)	(1.08)
1994 Dummy	-11.217	-11.067	-2.298	-11.486
	(1.81)	(6.32)**	(0.36)	(3.59)**
Chemicals Dummy	0.275	-0.002	-0.619	0.358
	(0.19)	(0.00)	(0.09)	(0.42)
Drugs and Medical Dummy	0.048	-0.345	-0.325	-0.015
	(0.03)	(0.11)	(0.04)	(0.02)
Electronics Dummy	-11.034	-11.488	-11.108	-11.223
	(6.75)**	(3.53)**	(1.38)	(3.23)**
Mechanical Dummy	-0.502	-0.663	0.035	-0.793
	(0.31)	(0.20)	(0.00)	(0.63)
Constant	11.179	11.341	15.260	11.471
	(6.45)**	(2.99)**	(1.36)	(6.66)**
Observations	268	243	55	277

Absolute value of *t*-statistics in parentheses
* significant at 5% level; ** significant at 1% level

Table 8: Tobit Estimation for Pre and Post-License Citations, Licensed Technologies Only

	Dependent Variable: Log(1 + License Revenues)		
	Pooled Model	Pre- and Post- License Citations	Licensee and Non-Licensee Citations
Total Citations	0.024		
	(0.55)		
Licensee Citations		-0.440	
		(1.45)	
Non-Licensee Citations		0.059	
		(1.22)	
Post-License Citations			0.012
			(0.24)
Pre-License Citations			0.096
			(0.78)
1982 Dummy	-9.178	-34.173	-9.632
	(1.22)	(1.92)	(1.28)
1983 Dummy	-5.810	-31.650	-6.275
	(0.68)	(1.69)	(0.74)
1984 Dummy	-2.759	-28.621	-2.918
	(0.37)	(1.57)	(0.39)
1985 Dummy	-8.685	-34.444	-9.170
	(1.13)	(1.88)	(1.19)
1986 Dummy	-14.006	-39.870	-14.503
	(1.61)	(2.12)*	(1.67)
1987 Dummy	-3.099	-28.507	-3.530
	(0.37)	(1.55)	(0.42)
1988 Dummy	-8.053	-35.143	-8.857
	(0.99)	(1.82)	(1.07)
1989 Dummy	-6.367	-32.806	-6.918
	(0.73)	(1.71)	(0.79)
1990 Dummy	-7.974	-33.799	-8.391
	(1.04)	(1.84)	(1.10)
1991 Dummy	-3.453	-28.929	-3.516
	(0.40)	(1.56)	(0.41)
1992 Dummy	-5.735	-31.639	-6.142
	(0.69)	(1.70)	(0.68)
1992 Dummy	-5.288	-30.883	-5.692
	(0.63)	(1.67)	(0.74)
1994 Dummy	-3.199	-28.831	-3.514
	(0.39)	(1.56)	(0.43)
Chemicals Dummy	-0.993	-2.441	-1.229
	(0.16)	(0.41)	(0.20)
Drugs and Medical Dummy	-2.592	-3.873	-2.603
	(0.44)	(0.66)	(0.44)
Electronics Dummy	-10.546	-11.868	-10.764
	(1.80)	(2.04)*	(1.84)
Mechanical Dummy	-1.962	-1.995	-3.451
	(0.24)	(0.25)	(0.41)
Constant	18.618	46.089	19.160
	(1.98)	(2.30)*	(2.04)*
Observations	56	56	56

Absolute value of *t*-statistics in parentheses
* significant at 5% level; ** significant at 1% level

Table 9: CLAD Estimation for Pre and Post-License Citations, Licensed Technologies Only

	Dependent Variable: Log(1 + License Revenues)		
	Pooled Model	Pre- and Post- License Citations	Licensee and Non-Licensee Citations
Total Citations	0.032		
	(0.41)		
Licensee Citations		0.278	
		(0.59)	
Non-Licensee Citations		0.084	
		(0.98)	
Post-License Citations			0.025
			(0.29)
Pre-License Citations			0.070
			(0.16)
1982 Dummy	-1.156	17.960	-1.907
	(0.10)	(0.76)	(0.22)
1983 Dummy	-4.291	15.464	-4.550
	(0.48)	(0.64)	(0.52)
1984 Dummy	-1.665	14.027	-1.803
	(0.26)	(0.62)	(0.24)
1985 Dummy	-5.567	13.132	-6.873
	(0.67)	(0.55)	(0.68)
1986 Dummy	-6.454	5.104	-15.267
	(1.31)	(0.20)	(1.52)
1987 Dummy	-0.716	19.038	-0.975
	(0.09)	(0.74)	(0.10)
1988 Dummy	-4.625	15.340	-4.912
	(0.52)	(0.61)	(0.46)
1989 Dummy	-2.642	16.534	-3.099
	(0.39)	(0.63)	(0.35)
1990 Dummy	-4.688	10.847	-5.053
	(0.43)	(0.49)	(0.59)
1991 Dummy	-3.458	15.928	-3.669
	(0.44)	(0.66)	(0.40)
1992 Dummy	-4.178	16.885	-3.555
	(0.58)	(0.65)	(0.31)
1992 Dummy	-3.262	16.467	-4.472
	(0.44)	(0.70)	(0.48)
1994 Dummy	-2.298	16.413	-2.371
	(0.32)	(0.66)	(0.26)
Chemicals Dummy	-0.619	-4.988	-0.760
	(0.08)	(0.73)	(0.13)
Drugs and Medical Dummy	-0.325	-3.609	-0.451
	(0.04)	(0.44)	(0.07)
Electronics Dummy	-11.108	-15.141	-11.243
	(1.16)	(1.70)	(1.50)
Mechanical Dummy	0.035	-9.039	-1.783
	(0.00)	(0.74)	(0.11)
Constant	15.260	-1.578	15.693
	(1.19)	(0.06)	(1.66)
Observations	54	50	55

Absolute value of *t*-statistics in parentheses
* significant at 5% level; ** significant at 1% level

Table 10: Tobit and CLAD Estimates, Including Year-Citation Interaction Terms, UC and Columbia, Licensed Technologies Only

	Tobit	CLAD
Citations*1980 Dummy	0.051	0.054
	(1.20)	0.009
Citations*1981 Dummy	0.073	0.098
	(0.49)	(0.13)
Citations*1982 Dummy	-0.059	0.038
	(0.89)	(0.45)
Citations*1983 Dummy	-0.077	-0.005
	(0.75)	(0.02)
Citations*1984 Dummy	-0.009	0.050
	(0.13)	(1.34)
Citations*1985 Dummy	0.003	0.018
	(0.05)	(0.51)
Citations*1986 Dummy	0.162	0.097
	(1.34)	(0.56)
Citations*1987 Dummy	0.029	-0.090
	(0.22)	(1.24)
Citations*1988 Dummy	0.026	0.042
	(0.49)	(1.27)
Citations*1989 Dummy	-0.146	0.062
	(1.21)	(0.73)
Citations*1990 Dummy	0.096	(0.45)
	(1.61)	(0.26)
Citations*1991 Dummy	0.128	0.065
	(0.94)	(0.23)
Citations*1992 Dummy	0.902	1.082
	(2.21)*	(1.59)
Citations*1993 Dummy	0.186	0.337
	(0.53)	(0.31)
Citations*1994 Dummy	0.423	0.428
	(2.65)**	(1.47)
1981 Dummy	-3.761	-2.012
	(1.06)	(0.43)
1982 Dummy	-1.181	-2.440
	(0.41)	(0.83)
1983 Dummy	-4.059	-1.355
	(1.38)	(0.36)
1984 Dummy	0.024	-0.701
	(0.01)	(0.23)
1985 Dummy	-3.483	-0.219
	(1.35)	(0.07)
1986 Dummy	-8.609	-2.604
	(2.80)**	(0.50)
1987 Dummy	-3.326	0.249
	(1.23)	(0.08)
1988 Dummy	-2.135	-0.962
	(0.86)	(0.32)
1989 Dummy	-1.602	-1.225
	(0.63)	(0.39)
1990 Dummy	-4.586	-0.147
	(1.88)	(0.05)
1991 Dummy	-6.997	-2.088
	(2.70)**	(0.63)
1992 Dummy	-12.291	-11.508
	(4.04)**	(1.89)
1993 Dummy	-9.075	-2.840
	(3.59)**	(0.52)

1994 Dummy	-14.069	-12.365
	(5.41)**	(2.72)**
Chemicals Dummy	-0.061	0.981
	(0.03)	(0.54)
Drugs and Medical Dummy	-1.396	0.567
	(0.65)	(0.31)
Electronics Dummy	-13.317	-10.831
	(5.71)**	(6.18)**
Mechanical Dummy	-2.780	-0.345
	(1.03)	(0.12)
Constant	14.234	10.941
	(4.88)**	(3.03)**
Observations	357	357

Absolute value of *t*-statistics in parentheses
* significant at 5% level; ** significant at 1% level

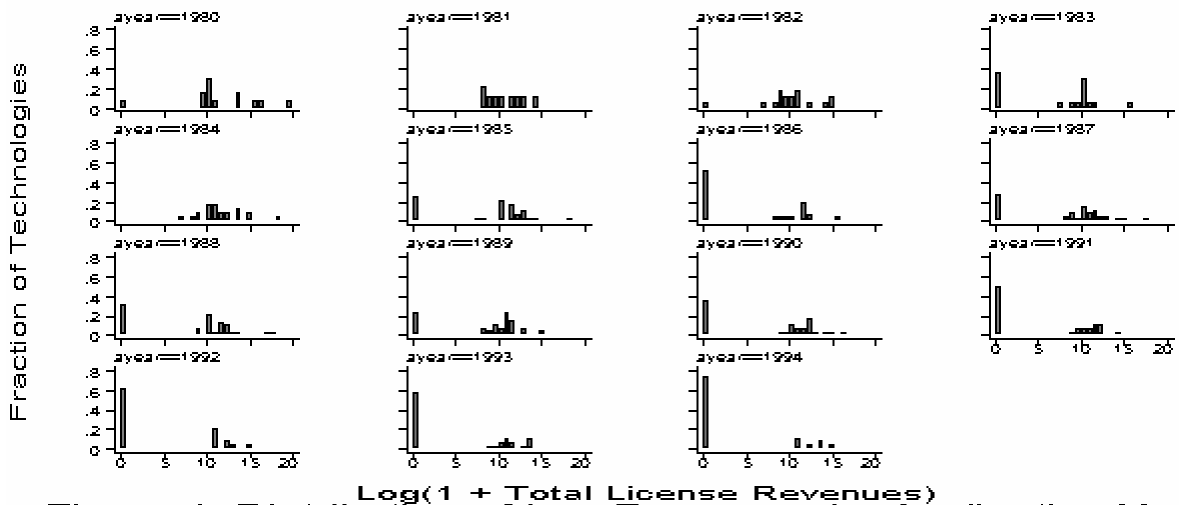


Figure 1: Distribution of Log Revenues by Application Year

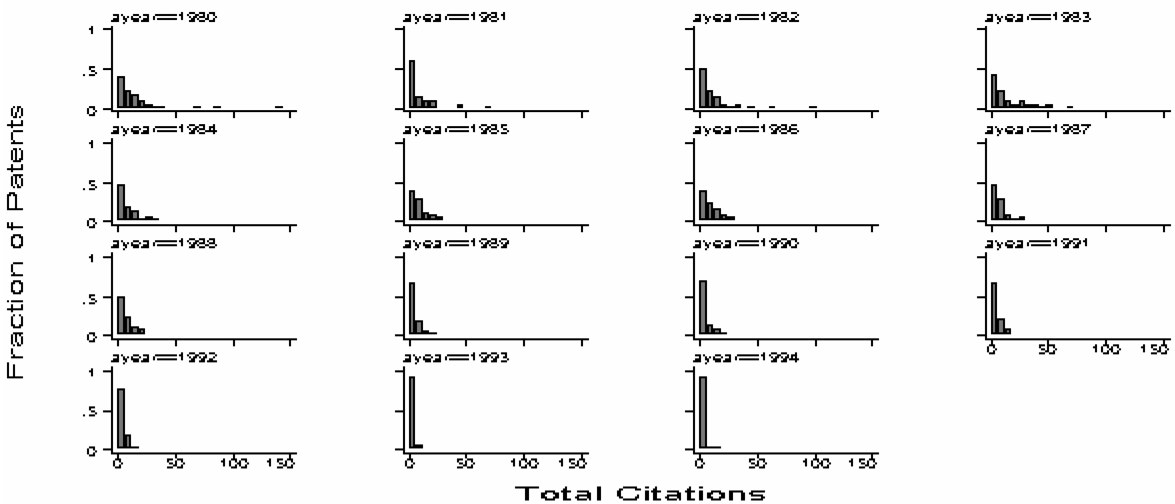


Figure 2: Distribution of Citations by Application Year

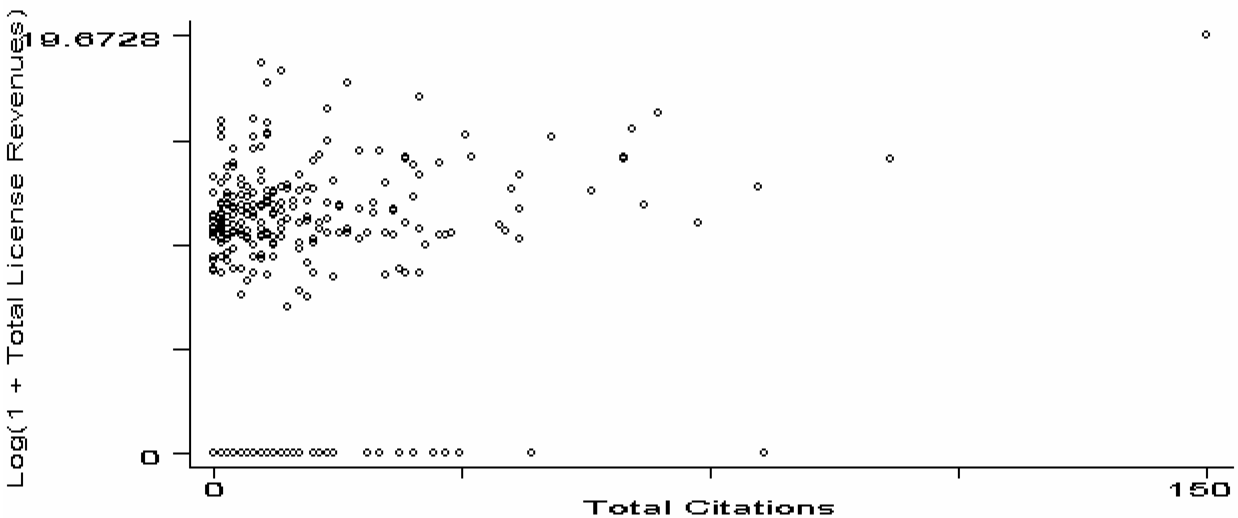


Figure 3: Scatterplot of Revenues versus Citations

Figure 4: Predicted Marginal Effects, from Probit Model

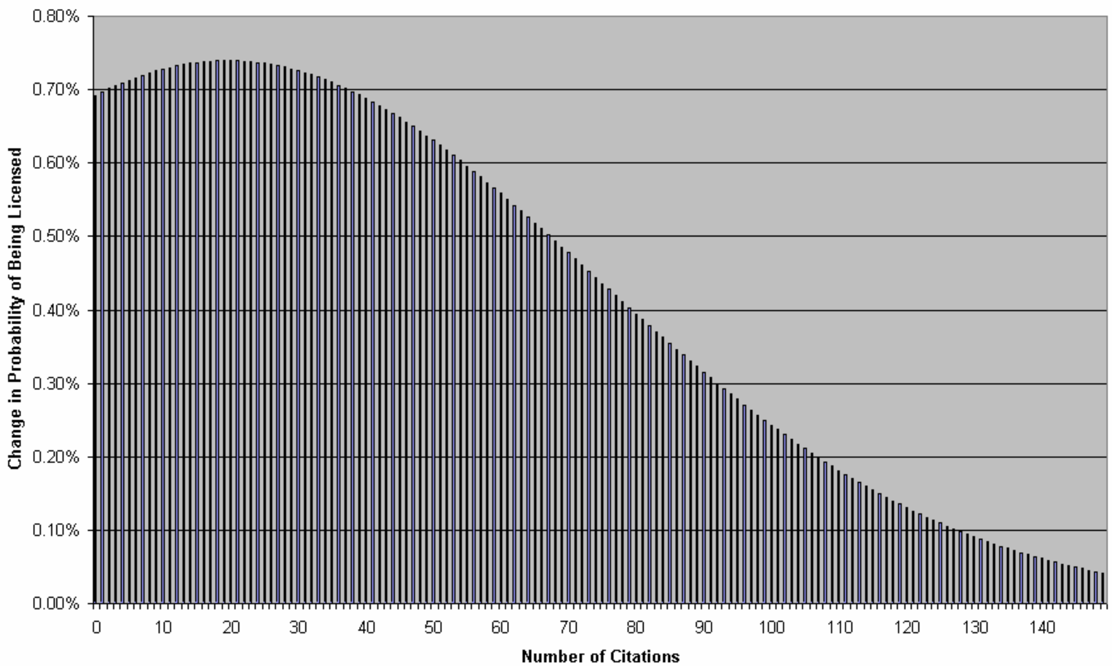
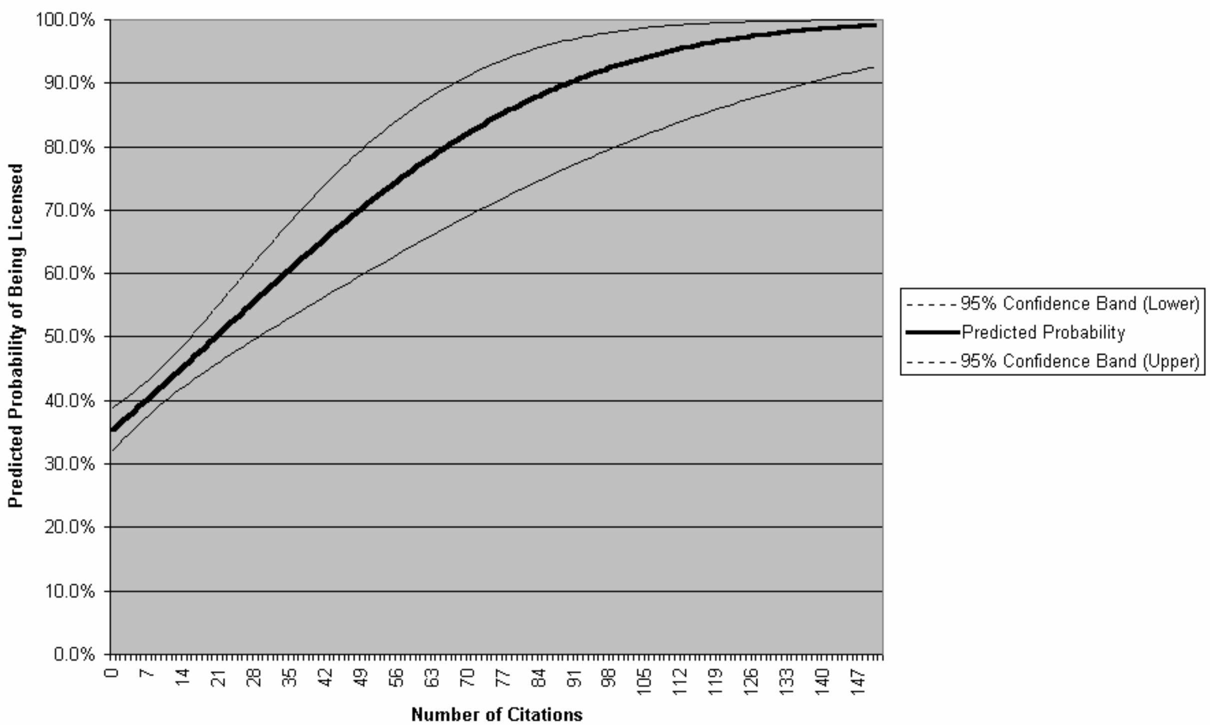


Figure 5: Predicted Probabilities of Being Licensed, from Probit Model



References

1983-2000. *Directory of Corporate Affiliations*. Skokie, Ill.: National Register Pub. Co.

Adelman, Martin, Randall Rader, John Thomas, and Harold Wegner. 1998. *Cases and Materials on Patent Law*. St. Paul: West Group.

Arabmazar, Abbas, and Peter Schmidt. 1982. "An Investigation of the Robustness of the Tobit Estimator to Non-Normality." *Econometrica* 50:1055-63.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention." in *The Rate and Direction of Inventive Activity*, edited by Richard Nelson. Princeton: Princeton University Press.

Austin, David H. 1993. "An Event-Study Approach to Measuring Innovative Output: The Case of Biotechnology." *American Economic Review* 83:253-58.

Austin, David H. 1994. "Patent Citations and Appropriability," Resources for the Future.

Buchinsky, Moshe. 1994. "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression." *Econometrica* 62:405-58.

Buderi, Robert. 1999. "In Search of Innovation." *Technology Review*.

Caballero, Ricardo J., and Adam B. Jaffe. 1993. "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth." : MIT and NBER; Harvard U and NBER.

Caudill, S., and J. Jackson. 1989. "Measuring Marginal Effects in Limited Dependent Variable Models." *The Statistician* 38:203-6.

Cole, Jonathan R. 2000. "A Short History Of The Use Of Citations As A Measure Of The Impact Of Scientific And Scholarly Work." Pp. 281-300 in *The Web of Knowledge : A Festschrift in Honor of Eugene Garfield*, edited by Blaise Cronin and Helen Atkins. Medford, N.J.: Information Today.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, MD: Published for the World Bank by Johns Hopkins University Press.

Deng, Zhen, Baruch Lev, and Francis Narin. 1999. "Science & Technology as Predictors of Stock Performance." *Financial Analysts Journal* 55.

Garfield, Eugene. 1957. "Breaking the Subject Index Barrier--A Citation Index for Chemical Patents." *Journal of the Patent Office Society* 34:583-593.

Griliches, Zvi. 1986. "Economic Data Issues." Pp. ages 1466-1514 in *Handbook of Econometrics. Volume 3*, edited by Zvi Griliches and Michael- D. Intriligator. Handbooks in Economics series, Book 2 (Amsterdam; Oxford and Tokyo: North-Holland) distributed in the U.S. and Canada by Elsevier Science.

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

Griliches, Zvi. 1992. "The Search for R&D Spillovers." *Scandinavian Journal of Economics* 94:S29-47.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. 2000. "Market Value and Patent Citations: A First Look." *NBER Working Paper W7741*.

Harhoff, Dietmar, and et al. 1999. "Citation Frequency and the Value of Patented Inventions." *Review of Economics and Statistics* 81:511-15.

Harhoff, Dietmer, Frederic Scherer, and Katrin Vopel. 1999. "Citations, Family Size, and the Value of Patent Rights." *Mimeo*.
Harmon, Robert L. 1999. *Patents and the Federal Circuit*. Washington, D.C.: The Bureau of National Affairs, Inc.
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* 80:119-27.
Jaffe, Adam. 1998. "Patents, Patent Citations, and the Dynamics of Technological Change." *NBER Reporter* 1998.
Jaffe, Adam, Rebecca Henderson, and Manuel Trajtenberg. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108.
Jaffe, Adam, and Josh Lerner. 1999. "Privatizing R&D: Patent Policy and the Commercialization of National Laboratory Technologies." *NBER Working Paper W7064*.
Jaffe, Adam, Manuel Trajtenberg, and Michael Fogarty. 2000. "The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees." *NBER Working Paper W7631*.
Jaffe, Adam B., Michael S. Fogarty, and Bruce A. Banks. 1998. "Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation." *Journal of Industrial Economics* 46:183-205.
Jaffe, Adam B., and Manuel Trajtenberg. 1999. "International Knowledge Flows: Evidence from Patent Citations." *Economics of Innovation and New Technology* 8:105-36.
Jaffe, Adam B., and Manuel Trajtenberg. 2000. Patents, Citations & Innovations: A Window on the Knowledge Economy (Cambridge, MA: MIT Press).
Jefferson, Thomas. 1984. *Writings of Thomas Jefferson*. New York: The Library of America.
Jensen, Richard, and Marie Thursby. 1998. "Proofs and Prototypes for Sale: The Tale of University Licensing." *American Economic Review* .
Klepper, Steven, and Edward E. Leamer. 1984. "Consistent Sets of Estimates for Regressions with Errors in All Variables." *Econometrica* 52:163-83.
Koenker, Roger, and Gilbert Bassett, Jr. 1982. "Robust Tests for Heteroscedasticity Based on Regression Quantiles." *Econometrica* 50:43-61.
Koenker, Roger W., and Gilbert Bassett, Jr. 1978. "Regression Quantiles." *Econometrica* 46:33-50.
Krasker, William S., and John W. Pratt. 1986. "Bounding the Effects of Proxy Variables on Regression Coefficients." *Econometrica* 54:641-55.
Krugman, Paul. 1991. *Geography and trade*. Gaston Eyskens Lecture Series, Cambridge, Mass. and London.
Lanjouw, Jean O., and Mark Schankerman. 1999. "The Quality of Ideas: Measuring Innovation with Multiple Indicators." *NBER Working Paper W7345*
Lanjouw, Jean O., and Mark Schankerman. 2000. "Characteristics of Patent Litigation: A Window on Competition." *Draft Manuscript* .
Maddala, G. S. 1983. *Limited-dependent and qualitative variables in econometrics*. New York: Cambridge University Press.
McDonald, John F., and Robert A. Moffitt. 1980. "The Uses of Tobit Analysis." *Review of Economics and Statistics* 62:318-21.

Melkers, Julia. 1993. "Bibliometrics as a Tool for Analysis of R&D Impacts." in *Evaluating R&D Impacts: Methods and Practice*, edited by Barry Bozeman and Julia Melkers. Boston: Kluwer.

Mogee, Mary Ellen, Richard G. Kolar, and Jonathan D. Putnam. 1997. "Patent Indicators of Global Technology Strategy." Reston, VA: Mogee Research & Analysis Associates.

Mowery, David C., and Bhaven N. Sampat. 2001. "Patenting and Licensing University Inventions: Lessons from the History of Research Corporation." *Industrial and Corporate Change*.

Mowery, David C., and Bhaven N. Sampat. 2001. "University Patents and Patent Policies: 1925-1980." *Industrial and Corporate Change*.

Mowery, David C., Bhaven N. Sampat, Richard R. Nelson, and Arvids A. Ziedonis. 2001. "The Growth of Patenting and Licensing By U.S. Universities: An Assessment of the Effects of the Bayh-Dole Act of 1980" *Research Policy* 30: 99-119.

Mowery, David C., Bhaven N. Sampat, and Arvids A. Ziedonis. 2001. "Learning to Patent? Institutional Experience and the Quality of University Patents" *Management Science* 48(1): 73-89.

Powell, James L. 1986. "Censored Regression Quantiles." *Journal of Econometrics* 32:143-55.

Rosenberg, Peter. 2000. *Patent Law Fundamentals*. St. Paul: West Group.

Sampat, Bhaven N., David C. Mowery, and Arvids A. Ziedonis. 2002. "Changes in University Patent Quality After The Bayh-Dole Act: A Re-Examination." *Draft Manuscript*.

Scherer, F.M., and Dietmer Harhoff. 2000. "Technology policy for a world of skew-distributed outcomes." *Research Policy* 29:559-566.

Schlicher, John. 1992. *Patent Law: Legal and Economic Principles*. New York: Clark, Boardman, and Callaghan.

Shane, Hilary, and Mark Klock. 1997. "The Relation between Patent Citations and Tobin's Q in the Semiconductor Industry." *Review of Quantitative Finance and Accounting* 9:131-46.

Sheldon, Jeffrey. 1998. *How to Write A Patent Application*. New York: Practicing Law Institute.

Thomas, John R. 2000. "A Proposal for Patent Bounties." *Unpublished Manuscript, George Washington University Law School*.

Trajtenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *Rand Journal of Economics* 21:172-87.

Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe. 1997. "University versus Corporate Patents: A Window on the Basicness of Invention." *Economics of Innovation and New Technology* 5:19-50.

United States Patent and Trademark Office. 2000. *Manual of Patent Examining Procedure*. Washington: U.S. Government Printing Office.