Exhibit 2027
Zynga, Inc. v. Personalized Media Communications, LLC
Case IPR2013-00164 (SCM)

Market value and patent citations

Bronwyn H. Hall*

Adam Jaffe**

Manuel Trajtenberg***

Abstract

This paper explores the usefulness of patent citations as a measure of the "importance" of a firm's patents, as indicated by the stock market valuation of the firm's intangible stock of knowledge. Using patents and citations for 1963-1999, we estimate Tobin's q equations on the ratios of R&D to assets stocks, patents to R&D, and citations to patents. We find that each ratio significantly impacts market value, with an extra citation per patent boosting market value by 3%. Further findings indicate that "unpredictable" citations have a stronger effect than the predictable portion, and that self-citations are more valuable than external citations.

JEL Classification: O31, O38

* UC Berkeley, NBER, and IFS UC Berkeley, NBER, and IFS; Department of Economics 549 Evans Hall, UC Berkeley, Berkeley, CA 9472—3880, USA; Tel. 1 510 642 3878; bhhall@econ.berkeley.edu

** Brandeis University and NBER; Brandeis University, Waltham MA 02554-9110, USA: 781-736-3451; ajaffe@brandeis.edu

*** Tel Aviv University, NBER, and CEPR; Eitan Berglas School of Economics, Tel Aviv University, Tel Aviv 69978, Israel; Tel. 972-3-640-9911; manuel@post.tau.ac.il

This paper follows closely in the steps of the late Zvi Griliches and owes to him the underlying vision, method, and pursuit of data. The data construction was partially supported by the National Science Foundation, via grants SBR-9413099 and SBR-9320973. We are extremely grateful to Meg Fernando of REI, Case Western Reserve University, for excellent assistance in matching the patent data to Compustat. We also acknowledge with gratitude the comments received at numerous seminars, and from two referees.

1. Introduction

It is widely understood that the R&D conducted by private firms is an investment activity, the output of which is an intangible asset that can be labeled as the firm's "knowledge stock." If this asset is known to contribute positively to the firm's future net cash flows, then the size of a firm's knowledge stock should be reflected in the observed market value of the firm. This implies that a firm's R&D investments should be capitalized in the firm's market value. Further, since the output of the R&D investment process is stochastic, some of the R&D will result in the creation of more valuable knowledge capital; if this success is observable, then it should be reflected in greater market value bang for the R&D buck.

Empirical testing of this formulation requires an observable proxy for R&D "success." There is a considerable literature using counts of firms' successful patent applications for this purpose. But the value of patent counts as a proxy for R&D success is severely limited by the very large variance in the significance or value of individual patents, rendering patent counts an extremely noisy indicator of R&D success. In this paper we utilize information on the number of subsequent citations received by a firm's patents to get a better measure of R&D success. Further, because citations arrive over time, and can be distinguished by the identity of the citing organization, we can distinguish the impact of a firm's patents on its market value according to the time path and source of subsequent citations.

This project was made possible by the recently completed creation of a comprehensive data file on patents and citations, comprising all US patents granted during the period 1963-1999 (three million patents), and all patent citations made during 1975-1999 (about 16 million citations), as described in Hall, Jaffe, and Trajtenberg (2001).[1] We construct on the basis of these data three measures of "knowledge stocks": the

[1] The complete data are available in the NBER site at http://www.nber.org/patents/, and also in a CD included with Jaffe and Trajtenberg (2002). For purposes of this paper, we actually used a previous version of the data that extends only until 1996.

traditional R&D and patent count stocks, and a citations stock. The last poses serious truncation problems, since citations to a given patent typically keep coming over long periods of time, but we only observe them until the last date of the available data; we apply correction methods developed elsewhere to deal with this and related problems. It is important to note that in this paper we look only at a simple "hedonic" (and hence snapshot-like) market value equation, and do not address the deeper dynamic forces at work, as discussed by Pakes (1985) – these will have to wait for future research.

We estimate Tobin's q "hedonic" equations on three complementary aspects of knowledge stocks: R&D "intensity" (the ratio of R&D stocks to the book value of assets), the patent yield of R&D (i.e., the ratio of patent count stocks to R&D stocks), and the average citations received by these patents (i.e., the ratio of citations to patent stocks). We find that each of these ratios has a statistically and economically significant impact on Tobin's q. This confirms that the market values R&D inputs, values R&D output as measured by patents, and further values "high-quality" R&D output as measured by citation intensity.

When we look in more detail at the aspects of citation patterns that are associated with higher market value, we find: (i) The value of high citation intensity is disproportionately concentrated in highly cited patents: firms having two to three times the median number of citations per patent display a 35% value premium, and those with 20 citations and more command a staggering 54% market value premium. (ii) There are wide differences across sectors in the impact of each knowledge stock ratio on market value. (iii) Market value premia associated with patent citations confirm the forward-looking nature of equity markets: at a given point in time, market value premia are associated with future citations rather than those that have been received in the past, and the portion of total lifetime citations that is unpredictable based on the citation history at a given moment has the largest impact. (iv) Self-citations (i.e., those coming from down-the-line patents owned by the same firm) are more valuable than citations coming from external patents, but this effect decreases with the size of patent portfolio held by the firm, as might be expected.

The paper is organized as follows: section 2 discusses the rationale for the use of patent and citations data in this sort of research, and reviews previous literature. The data are described in section 3, along with a discussion of truncation and its remedies. Section 4 deals with the specification of the market value equation, and the construction of citation stocks, including the partition into past-future and predictable-residual citation stocks. The empirical findings are presented in section 5: starting with a "horse race" between R&D, patents, and citations, we proceed to estimate the preferred specification that includes the three ratios, add industry effects, experiment with the various partitions of the citations stock, and finally look at the differential impact of self-citations. Section 6 concludes with ideas for further research.

2. Patents, citations, and market value: where do we stand?

Patents have long been recognized as a very rich data source for the study of innovation and technical change. Indeed, there are numerous advantages to the use of patent data: each patent contains highly detailed information on the innovation; patents display extremely wide coverage in terms of technologies, assignees, and geography; there are already millions of them (the flow being of over 150,000 US Patent and Trademark Office [USPTO] patent grants per year); the data contained in patents are supplied entirely on a voluntarily basis, etc. There are serious limitations as well, the most glaring being that not all innovations are patented, simply because not all inventions meet the patentability criteria, and because the inventor has to make a strategic decision to patent, as opposed to relying on secrecy or other means of appropriability.[2]

[2] Unfortunately, we have very little idea of the extent to which patents are representative of the wider universe of inventions, since there is no systematic data about inventions that are not patented (see, however, Crepon, Duguet, and Mairesse, 1998). This is an important, wide-open area for future research.

The large-scale use of patent data in economic research goes back to Scherer (1965), Schmookler (1966), and Griliches (1984).[3] One of the major limitations of these research programs, extremely valuable as they had been, was that they relied exclusively on patent counts as indicators of innovative output.[4] However, it has long been recognized that innovations vary enormously in their technological and economic “importance” or “value,” and that the distribution of such “values” is extremely skewed. Thus simple patent counts are inherently limited in the extent to which they can capture such heterogeneity (see Griliches, Pakes, and Hall, 1987). The line of research initiated by Pakes and Schankerman (1984) using patent renewal data clearly revealed these features of the patent data. Patent citations suggested themselves as a means to tackle such heterogeneity (Trajtenberg, 1990; Albert et al., 1991), as well as a way to trace spillovers (Jaffe, Trajtenberg, and Henderson, 1993). In order to understand the role that patent citations have come to play in this context, we have to look more in detail into the patent document as a legal entity and as an information source.

A patent awards to inventors the right to exclude others from the unauthorized use of the disclosed invention, for a predetermined period of time.[5] For a patent to be granted, the innovation must fulfill the following criteria: (i) it has to be ***novel*** in a legally defined sense[6]; (ii) it has to be ***non-obvious***, in that a skilled practitioner of the technology would not have known how to do it; and (iii) it must be ***useful***, meaning that it has potential commercial value. If a patent is granted, an extensive public document is created. The front page of a patent contains detailed information about the invention, the inventor, the

---

[3] The work of Schmookler involved assigning patent counts to industries, whereas Griliches' project entailed matching patents to a sample of Compustat firms. In both cases the resulting data used were yearly ***patent counts*** by industries or firms. Scherer's project involved the creation of a “technology flow matrix” by industry of origin and industries of use.

[4] Of course, that is the best they could do at the time, given computer and data resources available.

[5] Whether or not this right translates into market power depends upon a host of other factors, including the legal strength of these rights, the speed of technical advance, the ease of imitation, etc.

[6] In the US that means “first to invent,” whereas in Europe and Japan it means “first to file.”

assignee, and the technological antecedents of the invention, including citations to previous patents. These citations serve an important legal function, since they delimit the scope of the property rights awarded by the patent. Thus, if patent *B* cites patent *A*, it implies that patent *A* represents a piece of previously existing knowledge upon which patent *B* builds, and over which *B* cannot have a claim. The applicant has a legal duty to disclose any knowledge of the prior art (and thus the inventor's attorney typically plays an important role in deciding which patents to cite), but the decision regarding which citations to include ultimately rests with the patent examiner, who is supposed to be an expert in the area and hence able to identify relevant prior art that the applicant misses or conceals.[7]

Thus, patent citations presumably convey information on two major aspects of innovations.[8] The first is linkages between inventions, inventors, and assignees along time and space. In particular, patent citations enable the quantitative, detailed study of spillovers, along geographical, institutional, and related dimensions. The second is that citations may be used as indicators of the "importance" of individual patents, thus introducing a way of gauging the enormous heterogeneity in the "value" of patents.[9] In this paper we concentrate on the latter aspect, with only a passing reference to citations as indicators of spillovers when dealing with self-citations.

[7] "During the examination process, the examiner searches the pertinent portion of the 'classified' patent file. His purpose is to identify any prior disclosures of technology…which anticipate the claimed invention and preclude the issuance of a patent; which might be similar to the claimed invention and limit the scope of patent protection…; or which, generally, reveal the state of the technology to which the invention is directed….If such documents are found they are made known to the inventor, and are 'cited' in any patent which matures from the application…Thus, the number of times a patent document is cited may be a measure of its technological significance." (Office of Technology Assessment and Forecast, 1976, p. 167).

[8] Citations allow one also to probe into other aspects of innovations, such as their "originality," "generality," links to science, etc. – see Trajtenberg, Henderson, and Jaffe (1997).

[9] The two are of course related: one may deem more "important" those patents that generate more spillovers, and vice versa. Most research so far has treated these two aspects separately, but clearly there is room to aim for an integrative approach.

There are reasons to believe that citations convey not just technological but also economically significant information: Patented innovations are for the most part the result of costly R&D conducted by profit-seeking organizations; if firms invest in further developing an innovation disclosed in a previous patent, then the resulting (citing) patents presumably signify that the cited innovation is economically valuable. Moreover, citations typically keep coming over the long run,[10] giving plenty of time to dissipate the original uncertainty regarding both the technological viability and the commercial worth of the cited innovation. Thus, if we still observe citations years after the grant of the cited patent, it must be that the latter had indeed proven to be valuable.

A detailed survey of inventors provides some direct evidence on citations as indicative of the presumed links across innovations (Jaffe, Trajtenberg, and Fogarty, 2000). A set of "citing inventors" answered questions about their patented inventions, about the relationship of these to previous patents cited in theirs as well as to technologically similar "placebo" patents that were not actually cited. A second set of (matched) "cited inventors" answered similar questions regarding the *citing* patents. The results confirm that citations do contain significant information on knowledge flows, but with a substantial amount of noise. The answers revealed significant differences between the cited patents and the placebos as to whether the citing inventor had learned anything from the cited patent, and precisely how and what she learned from it. However, as many as half of all citations did not seem to correspond to any kind of knowledge flow, whereas one-quarter of them indicate a strong connection between citing and cited patents.

[10] The mean backward citation lag hovers around 15 years (depending on the cohort), the median at about 10, and 5% of citations go back 50 years and more. The forward lag is more difficult to characterize because of the inherent truncation, but looking at citations to the oldest cohort in the data, that of 1975, we see that even after 25 years citations keep coming at a non-declining rate (see Jaffe and Trajtenberg, 2002, Ch. 13).

There have been a small number of studies that attempted to validate the use of patent citations as indicators of economic impact or value. Trajtenberg (1990) related the flow of patents in computed tomography (CT) scanners, a major innovation in medical technology, to the estimated social surplus due to improvements in this technology.[11] Whereas simple patent counts showed no correlation with the estimated surplus, citation-weighted patent counts turned out to be highly correlated with it, thus providing first-time evidence to the effect that citations carry information on the value of patented innovations. Recent work by Lanjouw and Schankerman (2003) also uses citations, along with other measures such as number of claims and number of countries in which an invention is patented, as a proxy for patent "quality." They find that a composite measure has significant power in predicting which patents will be renewed and which will be litigated, thus inferring that that these indicators are indeed associated with the private value of patents. Harhoff et al. (1999) survey German patent holders of US patents that were also filed in Germany, asking them to estimate the price at which they would have been willing to sell the patent right three years after filing. They find that the estimated value is correlated with subsequent citations, and that the most highly cited patents are very valuable, with a single citation implying an average value of about $1 million. Giummo (2003) examines the royalties received by the inventor/patent holders at nine major German corporations under the German Employee Compensation Act and reaches similar conclusions.

There is a substantial literature relating the stock market value of firms to various measures of "knowledge capital," and in particular to R&D and patents, going back to the landmark research program initiated by Griliches and coworkers at the NBER.[12] Hall (2000) offers a recent survey of this line of

[11] Consumer surplus was derived from an estimated discrete choice model of demand for CT scanners, based on purchases of scanners by US hospitals. Innovation manifested itself in the sale of improved scanners over time, i.e., scanners having better characteristics (e.g., speed and resolution).

[12] See, among others, Griliches (1981), Pakes (1985), Jaffe (1986), Griliches, Pakes, and Hall (1987), Connolly and Hirschey (1988), Griliches, Hall, and Pakes (1991), Hall (1993a), Hall (1993b), and Blundell, Griffith, and van Reenen (1999).

work: the typical finding is that patent counts do not have as much explanatory power as R&D in a market value equation, but they do appear to add some information above and beyond R&D. A few papers have tried to incorporate patent citations as well, albeit in the context of small-scale studies: Shane (1993) finds that, for a small sample of semiconductor firms in 1977-1990, patents weighted by citations have more predictive power in a Tobin's q equation than simple patent counts, entering significantly even when R&D stock is included. Citations-weighted patents also turned out to be more highly correlated with R&D than simple patent counts, implying that firms invest more efforts into patented innovations that ultimately yield more citations. Finally, Austin (1993) finds that citation-weighted counts enter positively but not significantly in an event study of patent grants in the biotechnology industry.

3. Data

For the purposes of this project we have brought together two large datasets and linked them via an elaborate matching process: the first is all patents granted by the USPTO between 1965 and 1996, including their patent citations; the second is firm data drawn from Compustat, including market value, assets, and R&D expenditures. The matching of the two sets (by firm name) proved to be a formidable, large-scale task, that tied up a great deal of our research efforts for a long time: Assignees obtain patents under a variety of names (their own and those of their subsidiaries), and the USPTO does not keep a unique identifier for each patenting organization from year to year. In fact, the initial list of corporate assignees of the 1965-1995 patents included over 100,000 entries, which we sought to match to the names of the approximately 6,000 manufacturing firms on the Compustat files, and to about 30,000 of their subsidiaries (obtained from the *Who Owns Whom* directory), as of 1989.[13] In addition to firms patenting under a variety of names (in some cases for strategic purposes), the difficulties in matching are

[13] Since ownership patterns change over time, ideally one would like to match patents to firms at more than one point in time; however, the difficulties of the matching process made it impossible to aim for more than one match.

compounded by the fact that there are numerous spelling mistakes in the names, and a bewildering array of abbreviations. As shown in Hall, Jaffe, and Trajtenberg (2001), we nevertheless succeeded in matching over half a million patents, which represent 50-65% (depending on the year) of all patents of US origin that were assigned to corporations during the years 1965 to 1995.[14,15] Still, the results presented here should be viewed with some caution, since they might be affected by remaining matching errors and omissions.

The Compustat data comprise all publicly traded firms in the manufacturing sector (SIC 2000-3999) between 1976 and 1995. After dropping duplicate observations and partially owned subsidiaries, and cleaning on our key variables, we ended up with an unbalanced panel of 4,864 firms (approximately 1,700 per year). The firms are all publicly traded on the New York, American, and regional stock exchanges, or over-the-counter on NASDAQ. The main Compustat variables used here are the market value of the firm at the close of the year, the book value of the physical assets, and the book value of the R&D investment. The market value is defined as the sum of the common stock, the preferred stock,[16] the long-term debt adjusted for inflation, and the short-term debt net of assets. The book value is the sum of net plant and equipment, inventories, and investments in unconsolidated subsidiaries, intangibles, and

---

[14] That is, the 573,000 matched patents compose 50-65% of all assigned patents (about one-quarter don't have an assignee) granted to US corporate inventors. Since Compustat includes firms that are traded in the US stock market only, most US patents of foreign origin are obviously not matched. The percentage matched is rather high, considering that the matching was done only to ***manufacturing*** firms, and only to those ***listed in Compustat***.

[15] In order to ensure that we picked up all the important subsidiaries, we examined and sought to assign all unmatched patenting organizations that had more than 50 patents during the period. A spot check of firms in the semiconductor industry suggests that our total patent numbers are fairly accurate, except for some firms for which we found a 5-15% undercount, due primarily to changing ownership patterns after 1989 – see Hall and Ziedonis (2001).

[16] That is, the preferred dividends capitalized at the preferred dividend rate for medium risk companies given by Moodys.

others (all adjusted for inflation).[17] The R&D capital stock is constructed using a declining balance formula and the past history of R&D spending with a 15% depreciation rate (for details see Hall, 1990).

Using the patents and citation data matched to the Compustat firms, we constructed patent stocks and citation-weighted patent stocks, applying the same declining balance formula used for R&D (also with a depreciation rate of 15%). Our patent data go back to 1964, and the first year for which we used a patent stock variable in the pooled regressions was 1975, so the effect of the missing initial condition (i.e., patents prior to 1964) should be small for the patent variable. The fraction of firms in our sample reporting R&D expenditures each year hovers around 60-70%, and the fraction of firms with a positive patent stock lies in the same range.[18] The yearly fraction of firms with current patent applications is about 35-40%, the percentage dropping steeply by the end of period because of the application-grant lag.

Dealing with truncation

Patent data pose two types of truncation problems, one regarding patent counts, the other citation counts. The first stems from the fact that there is a significant lag between patent applications and patent grants (averaging lately about two years). Thus, as we approach the last year for which there are data available (e.g., 1995 in the data used here), we observe only a small fraction of the patents applied for that eventually will be granted.[19] As shown in Appendix A, correcting for this sort of truncation bias is

[17] These intangibles are normally the goodwill and excess of market over book value from acquisitions, and do not include the R&D investment of the current firm, although they may include some value for the results of R&D by firms that have been acquired by the current firm.

[18] Even though there is substantial overlap between firms reporting R&D and those with patent stocks, the two sets are not nested: 19% of the firms with R&D stocks have no patents, while 13% of the firms with patent stocks report no R&D.

[19] Of course, the difficulty stems from the fact that we do not observe patent applications (and even if we did, we would not know which of them would eventually be granted), and that we date patents by their application rather than by their grant year.

relatively straightforward, and essentially involves using the application-grant empirical distribution to compute "weight factors." Thus, and using the results reported there, a patent count for, say, 1994 would be adjusted upwards by a factor of 1.166, implying that about 17% of the patents applied for in 1994 are expected to be granted after 1995, the last year of the data.

Citation counts are inherently truncated, since patents keep receiving citations over long periods of time (in some cases even after 50 years), but we observe at best only the citations given up to the present, and more realistically only up to the last year of the available data. Moreover, patents applied for in different years suffer to different extents from this truncation bias in citations received, and hence their citation intensity is not comparable and cannot be aggregated. For recent patents the problem is obviously more acute, since we only observe the first few years of citations. Thus, a 1993 patent that received ten citations by 1996 (the end of our data) is likely to be a higher citation-intensity patent than a 1985 patent that received 11 citations within our data period. Furthermore, although our basic patent information begins in 1964, we only have data on the citations made by patents beginning in 1976. Hence patents granted before 1976 experience truncation at the *beginning* of their citation cycle.[20]

We address the problem of truncated citations by estimating the shape of the citation-lag distribution, i.e., the fraction of lifetime citations (defined as the 30 years after the grant date) that are received in each year after patent grant. We assume that this distribution is stationary and independent of overall citation intensity. Given this distribution, we can estimate the total citations of any patent for which we observe a portion of its citation life simply by dividing the observed citations by the fraction of the population distribution that lies in the time interval for which citations are observed.[21] In the case of patents for

[20] Thus, a 1964 patent that received ten citations between 1976 and 1996 is probably more citation-intensive than a 1976 patent that received 11 citations over that same period.

[21] The details of the estimation of the citation lag distribution and the derived adjustment to citation intensity are described in Hall, Jaffe, and Trajtenberg (2000), Appendix D, and further adjustment procedures are developed in Hall, Jaffe, and Trajtenberg (2001).

which we observe the prime citation years (roughly years 3-10 after grant), this should give relatively accurate estimates of lifetime citations. On the other hand, when we observe only the first few years after grant (which is the case for more recent patents), the estimates will be much more noisy. In particular, the estimate of lifetime citations for patents with no citations in their first few years will be exactly zero, despite the fact that some of those patents will be eventually cited. Because of the increasing imprecision in measuring cites per patent as we approach the end of our sample period, our pooled regressions focus first on the 1976-1992 period, and then on the subset of years between 1979 and 1988.[22]

A first look at the data

Table 1 shows the sample statistics for the main variables used in the analysis, for the sample of observations analyzed in Tables 3 through 6: as expected, both market and book value, and the various knowledge stocks (R&D, patents, and citations), are extremely skewed, with the means exceeding the median by over an order of magnitude. The ratios R&D/Assets and Citations/Patents are distributed much more symmetrically, reflecting systematic size effects; however, the patent yield (Patents/R&D) retains a high degree of skewness and displays a large variance, indicating a rather weak correlation between the two stocks. Both the dependent variable (market to book value) and the candidate regressors in the models to be estimated exhibit a non-negligible amount of within variation, suggesting that there is interesting "action" in both the cross-sectional and the temporal dimensions.

Figure 1 shows the total citation and patenting rates per real R&D spending in our sample. Patent counts are adjusted for the application-grant lag, and citation counts are shown both corrected and uncorrected:

[22] Another issue that arises in this context is that the number of citations *made* by each patent has been rising over time, suggesting a kind of "citation inflation" that renders each citation less significant in later years. In this paper we choose not to make any correction for the secular changes in citation rates, with the cost that our extrapolation attempts become somewhat inaccurate later in the sample. For a detailed discussion of this issue, and of econometric techniques to deal with it, see Hall, Jaffe, and Trajtenberg (2001).

clearly, correcting for truncation has a dramatic impact on the series, particularly for recent years. Although the earlier years (1975-1985) show a steady decline in patenting and citation-weighted patenting per R&D dollar, one can clearly see the marked increase in patenting that began in 1986-1987. However, the yield begins to decline in about 1993, mostly because of rapid increases in R&D during that period. The corrected patent citation yield also begins to increase in 1986-1987 but does not decline later quite as much as the patent yield, reflecting an increase in the number of citations *made* per patent in the early to mid-nineties.

The distribution of citations per patent is, as expected, extremely skewed: Fully one-quarter of the one million patents in our data have no citations, 150,000 have only one, 125,000 have two, and just four patents received more than 200 citations. Fitting a Pareto distribution to this curve yields a parameter of 1.8, which implies that the distribution has a mean but no variance. However, a Kolmogorov-Smirnov or other distributional test would easily reject that the data are actually Pareto. The most highly cited patent since 1976 is patent #4,440,871 assigned to Union Carbide Corporation in 1984, for synthesizing a novel class of crystalline microporous silicoaluminophosphates, widely used as catalysts in chemical reactions. This patent received 227 citations in our data (i.e., up to 1996), and a total of 349 up to July 2003.

4. Model specification: the market value equation

We use a specification of the firm-level market-value function that is predominant in the literature: an additively separable linear specification, as was used by Griliches (1981) and his various co-authors. A notable advantage of this specification is that it assumes that the marginal shadow value of the assets is equalized across firms. The model is given by,

(1) $$V_{it} = q_t(A_{it} + \gamma K_{it})^{\sigma}$$

where V_{it} denotes the market value of firm i at time t, A_{it} ordinary physical assets, and K_{it} the firm's knowledge assets. The parameter σ allows for non-constant scale effects in the value function. All variables are in nominal terms. Taking logarithms we obtain,

(2) $$\log V_{it} = \log q_t + \sigma \log A_{it} + \sigma \log\left(1 + \gamma (K_{it} / A_{it})\right)$$

The last term is typically approximated by $\gamma(K_{it} / A_{it})$, in spite of the fact that the approximation can be relatively inaccurate for *K/A* ratios of the magnitude that are now common (above 15%). In this formulation, γ measures the shadow value of knowledge assets relative to the tangible assets of the firm, and $\sigma\gamma$ measures their absolute value. If the value function exhibits constant returns to scale (as it does approximately in the cross section), then $\sigma = 1$; in that case *logA* can be moved to the left hand side of the equation, and the model estimated with the conventional Tobin's q as the dependent variable. The estimating equation thus becomes,[23]

(3) $$\log Q_{it} \equiv \log\left(\frac{V_{it}}{A_{it}}\right) = \log q_t + \log\left(1 + \gamma \frac{K_{it}}{A_{it}}\right) + \varepsilon_{it}$$

where Q_{it} denotes Tobin's q, and the intercept of the model can be interpreted as an estimate of the logarithmic average of Tobin's q for each year. Theory does not give much guidance for the specification of knowledge stocks in an equation such as (3), and in particular it is not clear how to incorporate R&D, patents, and citation-weighted patents as measures of *K*. The more fundamental

[23] For one possible rationalization of the error term in the equation, see Griliches (1981).

question is how does innovative activity translate into market value, and what aspects of the underlying process are captured by the empirical measures available.[24]

We think of the knowledge creation process as a continuum going from R&D to patents to citations, which involves the sequential revelation of information about the value to the firm of the innovations generated along the way. That is, R&D reveals the commitment of firm's resources to innovation, patents catalog the success in generating codifiable new knowledge that can in principle be appropriated by the firm, and citations indicate the extent to which those innovations turn out to be "important" and hence presumably more valuable to the firm.

Once R&D is observed, the market presumably knows how to price the ***expected value*** of the innovative stream that will result from it, including the expected number of patents and citations that will come further down the line. Of course, the actual results from an R&D program will deviate from expectations, with some yielding "dry holes" and others unanticipated discoveries. Thus, the additional informational value of patents once R&D has already been factored in must reside in the number of patents ***per dollar R&D***: if the yield of R&D in terms of patents is higher than average that may indicate that the R&D project succeeded beyond expectations, and conversely if the patent yield is low. Once again, though, we know that the (ex post) value of patents is extremely skewed,[25] and that the stream of citations received over time is correlated with both the private and the social value of patented innovations (e.g., Harhoff et al., 1999; Trajtenberg, 1990). Therefore, the informational value of

[24] The discussion hereafter builds upon the pioneering work of Griliches (1984) and Pakes (1985), and subsequent work as in Griliches, Pakes, and Hall (1987), and Griliches, Hall, and Pakes (1991). However, we do not attempt to dwell in depth on the issues addressed in this literature, but just to sketch the arguments that lie behind the empirical specification used here. Once again, the use of citations in market value equations is new; previous work included only R&D and patents.

[25] A significant fraction of patents appear to be valueless — see Pakes and Schankerman (1984), and Pakes (1986).

citations once patents have been factored in must lie in the extent to which the number of received citations ***per patent*** deviates from expectations. The equation to be estimated thus becomes,

$$(4) \qquad \log Q_{it} = \log q_t + \log\left(1 + \gamma_1 \frac{R\&D_{it}}{A_{it}} + \gamma_2 \frac{PAT_{it}}{R\&D_{it}} + \gamma_3 \frac{CITES_{it}}{PAT_{it}}\right) + \varepsilon_{it}$$

where *R&D, PAT,* and *CITES* stand for the stocks of (lagged) R&D, Patents, and Citations respectively.

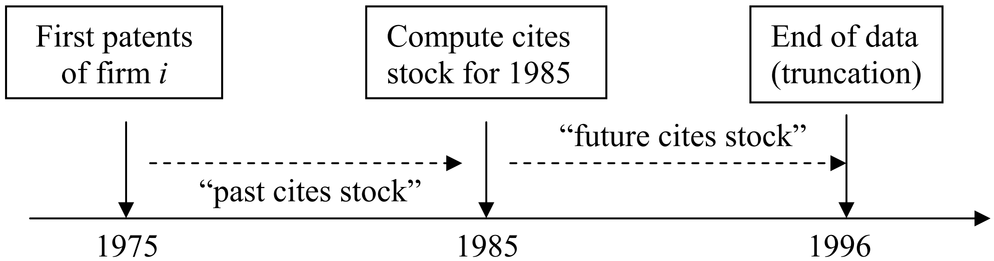
How to construct knowledge stocks?

The computation of R&D and patent stocks is relatively straightforward, essentially requiring just upfront assumptions as to the depreciation rate – here we shall just follow convention and resort to the traditional 15% depreciation rate, used in much of the literature.[26] The construction of citation stocks is inherently more complex, simply because citations are not a one-shot event at a point in time (as are R&D expenditures and patent applications), but keep coming over a long period of time into the future (some take 50 years and more).[27] In order to lay out the issues, consider the following time line[28]:

[26] Small departures from this rate do not make a difference to the results, but this is an issue that deserves some serious revisiting, in light of the much more detailed data that we have now at our disposal.

[27] Strictly speaking, even the granting of a patent is not the simple, one-time event described here; patent practice affords firms complex opportunities to divide and continue patent applications, so that in fact a given patent application can generate a stream of patent grants over time. We ignore this complication, treating each patent grant as a distinct event.

[28] As Quillen and Webster (2001) and Graham (2003) have shown, this time line is oversimplified due to the widespread use of patent application continuation to ensure early priority dates. In fact, about one-quarter of all patents (up to one half in the biotechnology area) have earlier priority dates but were delayed via continuation by the applicant (the typical delay is around half a year). In future work we plan to explore the consequences for the market value relationship of using the original application date for continued patents rather than the "official" application date recorded by the USPTO.



Consider a firm that started applying for patents in 1975, for which we want to compute its citation stock for every year from 1975 through the end of the period for which there are data available, 1996. In particular, consider the computation of the citation stock for one particular observation, that corresponding to 1985. The patents granted to this firm during the period 1975-1985 had received a certain number of citations up to 1985, which enter into what we shall call the “past citations stock,” i.e., the sum of citations known to have occurred *as of* 1985, properly discounted. However, since we are looking at it from the vantage point of a later year, we know already that these patents received further citations later on (i.e., during 1985-1996); these later citations can be used to compute the “future citations stock.” Provided that we formulate in a consistent manner the depreciation pattern (see below), we can then define the decomposition: (total citation stock)$_t$ = (past citations stock)$_t$ + (future citations stock)$_t$, $t = T_0, \ldots, 1996$, T_0 being the first year for which the firm applied for patents. We can use either of these stocks in the estimating equation (4), and thus investigate the timing of the revelation of information contained in citations, on the value of patented inventions. More formally, denote by $C(t, \tau)$ the number of citations received in year τ by patents applied for in year t; thus the total number of citations to year t patents observed till the end of the period is,

(5) $$C(t) = \sum_{\tau=t}^{1996} C(t,\tau)$$

Using the standard declining balance formula, the total citation stock observed in year t is,

(6) $$T\_CITES(t) = (1-\delta)T\_CITES(t-1) + C(t)$$

where δ stands for the (single) depreciation rate for the private value of patents. The "past" citations stock is computed as,

(7) $$P\_CITES(t) = (1-\delta)P\_CITES(t-1) + \sum_{s=0}^{T_0 - t}(1-\delta)^s C(t-s,t)$$

Note that in this formulation, patented innovations are assumed to have a value at the time they are applied for that is commensurate with the number of future citations, but we don't learn about such value until those future citations are received. Thus we depreciate citations as of the date when the receiving patents were applied for, but starting only when the citations are received.[29] The future stock is computed as,

(8) $$F\_CITES(t) = (1-\delta)F\_CITES(t-1) + \sum_{v=1}^{1996-t}\sum_{s=0}^{T_0 - t}(1-\delta)^s C(t-s,t+v)$$

To clarify, suppose for example that two citations to a 1980 patent are received in 1990, and that we compute *F_CITES(1985).* The second right-hand term in (8) will then be (for δ=*.15*): $2 \times 0.85^5 = 0.44$. That is, although these two citations came "late," their value as of 1985 is determined by the fact that these citations refer to an earlier patent (applied for in 1980), which value has depreciated substantially

[29] An alternative formulation would be that the patent becomes more valuable each time a citation is made, and hence starts depreciating citations as of the time the citations are made.

from 1980 to 1985.[30] In other words, the value of a patent at the moment it issues is proxied by the total number of citations that it will ever receive. This value depreciates as time passes; *T_CITES(t)* tracks this declining overall value. Distinct from this process of depreciation in "ultimate" value is the process of information revelation. At any point in time *t*, the (depreciated) value indicated by *T_CITES(t)* can be decomposed into the part that corresponds to citations that arrived before *t* (*P_CITES(t)*) and the part associated with citations that had not yet arrived as of t (*F_CITES(t)*).

The above split between past and future citations ignores the fact that past citations may be a good predictor of future ones, in which case the information available at time t from observing citations accrued until then would be not just the actual number of citations received, but the expected (total) number of citations that can be forecasted on that basis. Thus a further interesting decomposition of the citations stock would be,

$$(9) \qquad T\_CITES(t) = E\left[T\_CITES(t) \mid P\_CITES(t)\right] + R\_CITES(t)$$

where *R_CITES* stands for the "residual" citations stock, defined as the difference between the actual and the predicted stock on the basis of past citations. When estimating the market value equation we shall experiment also with this decomposition.

5. Estimating market value as a function of knowledge stocks

We estimate the market value equation (as in equation {4) above and variants of it) using nonlinear least squares, including year dummies, and in some specifications also industry fixed effects. We have chosen

[30] Note that the value of these two citations in the computation of *F_CITES(1985)* is independent of the year in which they are received.

not to include fixed firm effects for a number of reasons.[31] First, our interest centers on the market's valuation of R&D strategies across firms. Because R&D tends to change slowly over time, a firm's R&D intensity is highly correlated with its individual effect; in fact, it is an important component of what creates differences across firms, and thus removing them would entail an over-correction. A second but related problem is that our right-hand side variables are predetermined rather than exogenous and our panels are relatively short, requiring in principle the use of differenced estimators rather than within estimation to achieve consistency. But in the case of slowly changing right-hand side variables such as ours here, even small amounts of measurement error lead to substantial downward bias in the coefficient estimates when differenced estimators are used (Griliches and Hausman 1986). Finally, firms tend to change their strategies in response to market conditions and the assumption that differences across them are "fixed" or permanent is not a particularly good one.

First-cut estimation: "horse race" between R&D, patents, and citations

As a first pass at the data, and in order to make it comparable to previous studies, we estimate equation (3) with "*K*" defined either as the R&D stock, the patent stock, or the citations stock, for two subperiods, 1976-1984, and 1985-1992. Table 2 shows the results for two samples: all firms, and those with non-zero patents. Comparing the R^2s of the alternative specifications reveals that, as in previous studies, R&D stocks are more tightly correlated with market value than patents; the novel result here is that, even though R&D comes on top also vis-à-vis the citations stock, the latter fares better than the patents stock.

Interestingly, the fit sharply deteriorates when going from the first to the second subperiod, significantly more so for the equation with R&D in it (see Hall 1993a,b). In fact, we run also year-by-year regressions on each of the alternative stocks, and find that while the R^2s for all three stocks decline throughout the

[31] For a specification of the market value-R&D relationship that does incorporate firm effects see Blundell, Griffith, and Van Reenen (1999).

1980s, they tend to converge: by the mid-1980s citations stocks, and somewhat less so patent stocks, have as much explanatory power as R&D.[32] It is quite likely that these findings reflect the sharp changes in the patenting behavior of US firms that occurred in the early and mid-1980s, in particular the strengthening of patent rights, and the start of the spectacular rise in the rate of patenting (see Kortum and Lerner, 1998).

The coefficients of the alternative stocks are not directly comparable, since they are not in the same units: the coefficient of R&D/Assets is in dollars (of market value) per dollar (of R&D), whereas the coefficient of Patents/Assets is in terms of dollars per patent, and that of Citations/Assets is in dollars per citation. We thus normalize the coefficients of total R&D, of patents, and of citations by multiplying them by the ratio of total R&D, total patent, and total citation stocks respectively, to the deflated R&D stock for all firms.[33] Thus, for example, the normalized coefficient for R&D/Assets of 1.102 that appears in Table 2 (for the sample of all firms, first period) can be compared to the normalized coefficient of .253 = .607 x .417 for patents, and to .292 for citations. Once again, the marginal shadow value of R&D is much larger than that of patents or of citations, but the differences shrink in the second period.

In view of the truncation problems at both ends of our data period, and given that the shadow values of our measures appear to change over time, from now on we focus on the ten-year period in the middle of our sample, 1979-1988, when the data are the most complete, and the valuation coefficients rather stable.[34] We also confine the sample to observations on firms that had obtained at least one patent during

---

[32] Similar results are obtained for the coefficients themselves, with those for patents declining much less than those for R&D. If we take these findings at face value, they imply that firms have expanded their R&D spending to the point where the return on the marginal dollar has fallen, whereas the marginal return to obtaining an additional patent has remained roughly constant.

[33] We did not use the average (or median) of these ratios over firms because of the presence of many zeros, and the skewness of both the patents and the citations distributions; yet the ratio of the totals, as used here, is not very reliable either, and thus the normalized coefficients should be taken with (more than) a grain of salt.

[34] Based on yearly regressions of the model – see Hall, Jaffe, and Trajtenberg (2001), Figures 5a and 5b.

the period 1975-1988 (we refer to them as "patenting firms"), and hence for which we could compute a patent stock (and in principle also a citation stock) for the sample period.

Estimating the full model

We now turn to the estimation of the full model, as specified in equation (4): market value is assumed to depend not only on the R&D intensity of the firm (i.e., the ratio of the R&D stock to assets), but also on the patent yield (the ratio of the patent stock to R&D stock), and on the average "importance" of those patents as manifested in the ratio of citations to patents. As already mentioned, R&D, patents, and citations are seen as cascading indicators of the value of innovations, each adding further information on top of what could be predicted on the basis of the previous indicator. Thus, for a given level of R&D spending, firms that manage to patent more will presumably have higher market valuations, and similarly for firms with patent portfolios that receive on average more citations per patent.

Table 3 presents the results for alternative specifications of equation (4): columns (1) and (2) display the baseline estimates, (3) investigates more in detail the impact of the number of citations, and (4) and (5) experiment with the decomposition of citations into past and future, as well as into predicted and unpredicted. Each of the indicators shows a strong and highly significant impact on market value, and a substantial contribution to the overall fit of the estimated model. Thus, it is clear that each of the measures – R&D/Assets, Patents/R&D and Citations/Patents – adds information on top of what could be inferred just from the others.[35] In order to assess their quantitative impact, we need to compute the semi-elasticities,[36]

$$(10) \qquad \frac{\partial \log Q}{\partial (R\&D/A)} = \hat{\gamma}_1 \left(1 + \hat{\gamma}_1 \frac{R\&D}{A} + \hat{\gamma}_2 \frac{PAT}{R\&D} + \hat{\gamma}_3 \frac{CITES}{PAT} \right)^{-1}$$

and similarly for *PAT/R&D* and for *CITES/PAT*. The distributions of the ratios of the various stocks that appear in the right hand side of (10) are very skewed, and hence we evaluate them both at the mean and at the median, and we also compute the ratio of the totals. Table 4 presents the results, using the estimates $\{\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3\}$ from Table 3, column (2). Thus, an increase of one percentage point in the R&D intensity of a firm (i.e., in the ratio R&D/Assets) leads to a similar increase in market value, i.e., about .8%; an extra patent per million $ of R&D boosts market value by about 2%, and an extra citation per patent by over 3%. These are very substantial effects, confirming the hypothesized importance of knowledge stocks for firms' value. The impact of citations per patent is particularly striking, consistent with the "million dollar" worth of citations reported by Harhoff et al. (1999).

How valuable are highly cited patents?

We have already alluded to the fact that the distribution of citations is very skewed, with about one-quarter of all patents getting none, and only a few dozen (out of millions) receiving 100 citations and

[35] Note also that the dummy for not doing R&D is large and very significant. This occurs because the patent yield variable (Patents/R&D) has been set to zero when the firm has no R&D stock. The interpretation is that the average market value effect of being a firm with patents that does no R&D is 6-7%. There are 2,934 observations with no R&D in the current year; 1,960 have no stock of R&D either.

[36] Note that this is a partial derivative, holding the other ratios constant, which is not a trivial matter given that R&D appears also in the denominator of the patents ratio, and patents in the denominator of the citations-to-patents ratio.

more. This suggests that the average effect that we get in a regression such as that in Table 3, column (2), may not reveal the full extent of the impact of the tail of the citations distribution. We thus break the Citations/Patents variable up into five groups: < 5, 5-6, 7-10 (6 is the median), 11-20, > 20, and include dummy variables for each (the first serves as the base category). As shown in Table 3, column (3), for firms with fewer than the median number of citations per patent, it makes no difference how far below the median they fall: firms with 5-6 citations per patent have no higher value than those with less than 5. However, firms that average more than the median number of citations per patent exhibit a very significant increase in market value: 10% higher if having 7-10 citations per patent, and 35% higher if having 2-3 times the median (11-20 citations per patent). The most dramatic effect is for the 143 firms (573 observations) with more than 20 cites per patent: the market value of these firms is a staggering 54% higher than that would be expected given their R&D capital and their patent stock.

These 145 firms with exceptionally high citation rates are concentrated in computing, office equipment, semiconductors, and electronics (66 firms), and in pharmaceuticals and medical instruments (41 firms).[37] They include both small firms (which have very few highly cited patents) and medium to large firms (such as Intel, Compaq Computer, Tandem Computer, Alza, and General Signal Corporation). It appears that the larger firms are primarily in the electronics sector, broadly defined, while those in the pharmaceutical sector that average a high citation rate are more likely to be smaller biotechnology firms.

The timing of citations

As mentioned in section 3, citations unfold over time, sometimes well into the future, and hence it is not clear to what extent the market anticipates these future citations, or rather reacts to them as they occur. In Table 3, column (4), we partition the stock of citations into "past" and "future" as shown in equations (7)

[37] Of the remaining 38 firms, 8 are in machinery, 7 in textiles and apparel, and the other 23 scattered in various sectors.

and (8): whereas the coefficient of the stock based on future citations is somewhat larger than the coefficient of the stock based on all citations, that of past citations is *negative* and marginally significant. Thus, it would seem that past citations quickly become "old news" and hence do not impact current market value (beyond R&D and patents); on the other hand, investors are able to forecast pretty well the expected value of patented innovations as it evolves over time, which is later corroborated by future citations.

However, the past and future citations stocks are in fact highly correlated measures of the same underlying quantity, and thus this finding does not necessarily imply that realized citations are worthless for forecasting the value of the knowledge assets associated with patents or the expected profit stream from those assets. Indeed, past citations could be used to forecast future ones, and therefore to forecast also the value associated with the underlying patented innovations. To explore this idea, we decompose the total citation stock into the part predicted by the past citation stock and the (orthogonal) residual, as shown in equation (9).[38] As can be seen in Table 3, column (5), the coefficient of the unexpected portion is much larger than that of the predictable one, but the latter is still significantly positive. Thus, past citations clearly help in forecasting future returns, yet there is new and highly significant information on the value of patented innovations that is revealed along the way that correlates tightly with future/unpredicted citations.

Industry effects

In Table 5 we include dummies for six sectors: Drugs and Medical Instrumentation (henceforth just "Drugs"); Chemicals; Computers and Communications (henceforth just "Computers"); Electrical; Metals and Machinery; and miscellaneous (low-tech industries), and interact them with the knowledge stock

[38] We also included a full set of time dummies in the conditioning set.

ratios. In column (2) we can see that there is a high premium to being in the Drugs or Computers sector, which comes mostly at the expense of the coefficient of R&D intensity, which plunges by a half. The full interactions in column (3) reveal wide differences across sectors in the effects of each knowledge stock ratio. In general, the differential importance of patent yield and of citations per patent rises, at the expense of R&D intensity. Thus, whereas in no sector the effect of R&D/Assets is much larger than the average effect picked up in the base specification displayed in column (1), the impact of patent yield for Drugs is three times the average effect (.10 versus .031), and that of Computers twice as high; similarly but not as pronounced, the impact of Citations/Patents for Drugs is over 50% higher than the average effect, while that for Computers is small, and lower than that for the other sectors except for the low-tech sector. This contrast is consistent with the differing roles played by patents in the two sectors: Drugs is characterized by discrete product technologies where patents serve their traditional role of exclusion, and some of them are therefore valuable on an individual basis, as measured by citations. Computers and Communications is a group of complex product industries where any particular product may rely on various technologies embodied in several patents held by different firms. In this industry patents are largely valued for negotiating cross-licensing agreements, so their individual quality is not as important, although having them is.[39]

The impact of self-citations

Citations to a given patent may come from any subsequent patent, including from patents assigned to the same firm as that holding the cited patent. These "self-citations" may differ from other citations in various ways[40]: first, self-citations cannot be regarded as representing spillovers as these are commonly defined,

[39] See Cohen, Nelson, and Walsh (2000) for further discussion of this contrast and Hall and Ziedonis (2001) for evidence on semiconductors.

[40] Indeed, self-citations have received careful attention in previous studies, and in many cases they have been treated differently than external citations – see, e.g., Jaffe, Trajtenberg, and Henderson (1993); Jaffe and Trajtenberg (2002).

since they occur ***within*** the same economic unit. Second, although it is the patent examiner who decides what citations are made, this decision is based in part on information provided by the applicant, and occurs in a process of negotiation with the applicant's attorneys. Hence it is possible that self-citations are affected to some degree by firms' differential knowledge and incentives with respect to internal versus external citations. Finally, and most importantly for our purposes here, self-citations may provide very different signals than external citations regarding the value *to the firm* of the subsequent down-the-line, technologically connected innovations.

In general, the process of cumulative or sequential innovation embodies a process in which successive inventors compete away each other's excess returns (see, e.g., Scotchmer, 1991). If we think of sequences of patents linked by citations within this paradigm of cumulative innovation, then firms citing their own patents may be a reflection of the cumulative nature of innovation and the "increasing returns" property of knowledge accumulation, particularly within a narrow field or technology trajectory. Self-citations would then suggest that the firm has a strong competitive position in that particular technology and is in a position to internalize some of the knowledge spillovers created by its own developments. This would imply both that the firm has lower costs because there is less need to acquire technology from others, and that it may be able to earn higher profits without risking rapid entry (since it controls a substantial stretch of the underlying technology). If so, the ***private value*** of self-citations would be greater than that of external citations, since the latter might be indicative inter alia of spillovers that are beneficial to others, and in particular to potential or actual competitors. On the other hand, if the share of self-citations reflects primarily the extent of "self-bias," then their informational content as to market value might be weaker.

These effects may change with firm size, or rather with the size of their patent portfolio: the share of self-citations should increase with the size of the portfolio, simply because the more patents a firm has, the higher the probability that a citation from a new patent it gets will be given to a patent it already has.[41] Thus firms with larger portfolios will exhibit for "mechanical" reasons a larger share of self-citations, regardless of whether these self-citations are indicative of the type of real phenomena mentioned before (e.g., internalization of spillovers leading to competitive advantage). In fact, the raw correlation between share of self-citations and the log of the patent portfolio size is about .22.[42] Thus, the presumed link between self-citations and market value may weaken with the size of the patent portfolio. Moreover, it may well be that the "self-bias" increases with size (e.g., because of the presence of more active legal departments in firms with large portfolios), and thus the confluence of both the statistical effect just mentioned and the increasing self-bias may altogether cancel the expected positive effect of self-citations for firms holding large portfolios.

In order to investigate this question, we add self-citation measures to our valuation equation in two ways: the share of self-citations out of total citations, and the ratio of the stock of self-citations to the ***patent*** stock. Since self-citations are already included in the citation stock, when we add the latter variable to the regression, its coefficient will represent the premium or discount associated with self-citations. The distribution of the share of self-citations is very skewed (towards zero), the median share is .06, the mean .09, and for 1.5% of the observations the share is higher than .50.

Columns (2) and (3) of Table 6 show the results of estimating our preferred specification with these variables included. Both forms of the self-cite variable have highly significant and positive coefficients in the market value equation. From column (2) we see that an increase of ten percentage points in the share

[41] Suppose there are *N* patents that could be cited, and that they are all equally likely to be cited by a new patent, i.e., the probability of citing any particular patent is simply *1/N*. If the firm owns *n* out of these *N* patents, then the probability that a citation will be a self-citation is *n/N*. Thus the share of self-citations should grow with *n*, the size of the patent portfolio.

[42] Regressing the share of self-citations on log(patents), with year dummies and a dummy for zero cites, renders a coefficient of .013 on log(patents). Two standard deviations over the mean of log(patents) brings the share of self-citations to .055, almost twice the median.

of self-citations is associated with a 2.2% increase in market value. The results in column (3) indicate that, if each of the patents held by the firm receives an additional cite from patents held by other entities, market value increases by 4.3%; however, if those extra citations are made by the firm itself, market value rises 10%. In column (4) we investigate the question of the size of the patent portfolio held by the firm and the impact of self-citations: as hypothesized, the value-relevance of self-cites declines with size. For firms holding an average-sized patent portfolio (about 200 patents), the effect of self-cites is that just described, and for those with smaller portfolios self-cites have even a more pronounced effect on market value. However, for firms with a portfolio of 1,000 patents, self-citations do not make a difference, and above that the impact on market value is even negative.

6. Concluding remarks

The results of the paper clearly show that patent citations contain significant information on the market value of firms, in addition to R&D and simple patents counts, thus enriching the toolkit available to economists in trying to tackle empirically the intangible assets, and in particular the "knowledge stock" of firms. In so doing, our findings help overcome the problem of the huge heterogeneity in the "importance" of patents that greatly undermined their use in the explanation of firm value or performance. We should note, however, that substantial time is needed after a patent is granted to accumulate significant information about its citations. This means that citations-based analysis will never be usable for evaluation of current or very recent innovations.

It remains true that patent-related measures cannot win a "horse race" with R&D as a determinant of market value, but this is hardly surprising[43]: Even if citations are a reasonably informative signal of success, this does not mean that they will be more correlated with value than R&D, because optimizing

[43] The discussion in this paragraph relies heavily on points made by Samuel Kortum in an early comment on a draft of this paper.

firms will increase their R&D in response to success. Our findings indicate that the marginal effect of additional citations per patent on market value is very high (controlling for R&D intensity and for patent yield): if the "quality" of patents of a firm increases so that on average these patents receive one additional citation, the market value of the firm would increase by 3%. This is a very large impact, but then one has to keep in mind that getting an additional citation per patent is very hard, considering that the mean number of citations per patent is just over three, and that the distribution is extremely skewed, with about one-quarter receiving none. The finding of a high quasi-elasticity of Tobin's q to citations per patent (of 3%) implies, reassuringly, that the Citations/Patents ratio is not just a statistically significant regressor, but one that associates a large component of market value with patented innovations. Breaking the number of citations into segments reveals that as we move up the distribution to, say, twice the median, the impact on market value grows very large – a premium of about 35%, which increases to 50% as a firm collects citations to the tune of three times the median and more.

We classify firms into six major sectors in order to explore industry effects, and find that there are indeed wide differences across them in the impact of each knowledge stock ratio on market value. Thus, the impact of patent yield for Drugs is three times the average, and that of Computers twice as high; similarly, the impact of Citations/Patents for Drugs is over 50% higher than the average effect, while that for Computers is small, and lower than all others except for the low-tech sector.

Exploring in more detail the unfolding over time of the information that citations provide, we find that market value is highly correlated with the portion of down-the-line citations that cannot be predicted on the basis of past citations, and furthermore, the coefficient on these unpredictable citations is four times larger than the coefficient on the predictable portion. This confirms what one might expect: that the market "already knows" more about the value of particular innovations than the econometrician can figure out simply by manipulating historical data on citations. Interestingly, the market's evaluation of given innovations is ultimately confirmed by the arrival of future "unexpected" citations, "unexpected" in the sense that they cannot be predicted on the basis of past *citations*. This intriguing result calls for further inquiry: how many years' worth of citations does one have to observe in order to know most of what citations will eventually reveal – is, say, ten years enough? What fraction of what one learns from the lifetime citations (in the sense of the correlation with market value), is known after, say, five years? Also, one could explore whether this result is driven by the tail of the distribution, which we know is associated with much of the value. In other words, to what extent is it possible to predict that a patent will ultimately get more than 20 citations based only on the first few years' citations? Is this difficulty of predicting the really big winners what makes the unpredictable portion of the citations total so important?[44]

A further interesting finding is that market value is positively correlated with the share of self-citations out of total citations to a firm's patents, but that such linkage weakens with the size of the firm's patent portfolio. The self-citation variable gives us a window into technological competition, in the sense that it may inform us about the extent to which firms have internalized knowledge spillovers, or the strength of their competitive position vis-à-vis other firms in their industry. It would be interesting to link this kind of result to work on the gap between the social and private rates of return to innovation. Presumably, the fact that citations from other firms are associated with less value than one's own citations is connected to the likelihood that other firms' citations are potentially associated with their capturing part of the social return to the cited firm's innovation. A broader framework in which citations are related to the value of the citing firm as well as the cited firm may be fruitful in this regard.

[44] It would also be useful to explore the use of a functional form or normalization that would allow the relative value of past and future citations to be compared more directly, rather than just asking which adds more to the R^2.

In addition to these variations on the themes already struck, there are other aspects of citation behavior that are likely to be value-relevant and hence worth exploring.[45] One is *generality*, defined as (1 minus) the Herfindahl-Hirschman index of concentration of citations over patent classes (see Trajtenberg, Henderson, and Jaffe, 1997). The idea is that if citations to a patent are spread over a larger number of technological fields, the cited patent is to be regarded as more "general," in that it presumably spilled over a wider range of fields. In terms of impacting the market value of firms, though, one could hypothesize the following: for firms that concentrate in narrow fields of activity, more generality is likely to be detrimental, since the firm will not be able to appropriate the spillovers to other fields. For conglomerates, the opposite may be true. Thus, for example, we could compute the average generality of patents for firm j in year t, and interact it with a dummy for whether or not the firm is a conglomerate. We may need to normalize generality as well, since the measure depends on the number of citations. This suggests both a conceptual difficulty in separating the effects of generality from citation intensity, and also a complex truncation problem in the generality measure itself. There is thus a plentiful research agenda ahead; we hope the findings of this paper will encourage its pursuit.

[45] Another issue that we have explored little in this paper but which deserves attention is the variation in measurement error across measures based on widely varying numbers of patents and citations. Because of the count nature of the underlying data, measures based on few citations or patents are inherently "noisier" than those based on a large number. See Hall (2000) for a discussion of this issue in the context of a concentration index based on patent counts.

Appendix A

Correcting for truncation of patent counts

Figure A.1 shows the average distribution of the lag between application and grant for all US patents issued during the past four decades. The distributions are quite similar across decades, although there does seem to be a net reduction in the lag between the 1960s and the later periods.[46] Except for the 1960s, 95% of patent applications that are eventually granted will be granted by year 3, and 99% by year 5. Our measure of patents in a year is the number applied for that are ultimately granted, so our goal is to adjust the granted-application count at both ends of our sample. The fact that the median lag is short means that in making our adjustment we cannot go back before an application date of 1964 (because we only have grants made in 1967 and later) and that we will not be able to adjust the patent counts beyond 1993 (because we only see about half the patents applied for in 1994 due to the grant lags). We describe these adjustments in more detail below.

At the beginning of the sample (1967) we observe some of the patents applied for in 1964-1966, but not the ones that were granted so quickly that their grant date is before 1967. This suggests that we might be able to correct our application counts for 1964-1966 (that is, fill in for lags 0 to 3) using weights drawn from the distribution for the 1960s. At the end of the sample, the opposite happens: there are patents applied for between about 1991 and 1996 that are still pending; some of them will be granted eventually, meaning that our counts of successful patent applications for those years are too small. Again, we can scale up the numbers we do have using the grant-lag distribution. Therefore, we will compute the following two adjustments to our patent counts:

[46] The data for the 1990s are based only on the first half of the period and therefore longer issue lags will be truncated, implying that the apparent shortening of the issue lag may be an artifact.

$$
\text{(A1)} \quad
\begin{aligned}
\tilde{P}_t &= \frac{P_t}{\sum_{s=1967-t}^{M} w_s} \qquad 1964 < t < 1967 \\
\tilde{P}_t &= \frac{P_t}{\sum_{s=0}^{1994-t} w_s} \qquad 1991 < t < 1994
\end{aligned}
$$

where P_t is the number of patent applications at time t, M is the maximum issue lag (assumed to be equal to about 10), and the weights w_s are weights constructed from the average lag distributions shown in Figure A.1. In Table A.1 we show the weighting factors we will use (the inverse of $\widetilde{P}_t$ above). Note the edge effects, which imply that the 1996 data will not be usable, and that the 1964 and 1995 data will have more variance due to estimation error.[47]

[47] For this reason, we make no attempt to use data later than 1993 in the body of the paper (although we do use that data to construct stocks of future citations). The 1964 data are used only to the extent that they enter into the stocks of patents and citations that we construct. The first stock in our regressions is dated around 1973, so the measurement error effect should be quite small (recall that the counts are being depreciated by 15% per year).

References

Albert, M.B., Avery, D., Narin, F., and McAllister, P. “Direct Validation of Citation Counts as Indicators of Industrially Important Patents.” *Research Policy*, Vol. 20 (1991), pp. 251-259.

Austin, D. “An Event Study Approach to Measuring Innovative Output: The Case of Biotechnology.” *American Economic Review*, Vol. 83 (1993), pp. 253-258.

Blundell, R., Griffith, R., and van Reenen, J. “Market Share, Market Value, and Innovation in a Panel of British Manufacturing Firms.” *Review of Economic Studies*, Vol. 66 (1999), pp. 529-554.

Cohen, W.M., Nelson, R.R., and Walsh, J.P. “Protecting Their Intellectual Assets: Appropriability Conditions and Why Firms Patent or Not?” Working Paper No. 7552, NBER, 2000.

Connolly, R.A. and Hirschey, M. “Market Value and Patents: A Bayesian Approach.” *Economics Letters,* Vol. 27 (1988), pp. 83-87.

Crepon, B., Duguet, E., and Mairesse, J. “Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level.” *Economics of Innovation and New Technology*, Vol. 7 (1998), pp. 115-156.

Giummo, J. *An Empirical Examination of the Value of Patented Inventions Using German Employee Inventors' Compensation Records.* Ph.D. dissertation, University of California at Berkeley, 2003.

Graham, S.J.H. *Capturing Value with Secrecy: Three Studies Exploring Firms' Use of Continuation Patenting in Appropriating Returns from Technological Innovation*. Ph.D. dissertation, University of California at Berkeley, 2003.

Griliches, Z. “Market Value, R&D, and Patents.” *Economic Letters,* Vol. 7 (1981), pp. 183-187.

—, ed. *R&D, Patents and Productivity*. Chicago: University of Chicago Press, 1984.

— and Hausman, J.A. “Errors in Variables in Panel Data.” *Journal of Econometrics*, Vol. 31 (1986), pp. 93-118.

—, Pakes, A., and Hall, B.H. “The Value of Patents as Indicators of Inventive Activity.” In P. Dasgupta and P. Stoneman, eds., *Economic Policy and Technological Performance,* Cambridge, England: Cambridge University Press, 1987.

—, —, and —. “R&D, Patents, and Market Value Revisited: Is There a Second (Technological Opportunity) Factor?” *Economics of Innovation and New Technology*, Vol. 1 (1991), pp. 183-202.

Hall, B. H., “The Manufacturing Sector Master File: 1959-1987.” NBER Working Paper No. 3366 (1990)

—. “The Stock Market Valuation of R&D Investment during the 1980s.” *American Economic Review*, Vol. 83 (1993a), pp. 259-264.

—. “Industrial Research during the 1980s: Did the Rate of Return Fall?” *Brookings Papers on Economic Activity Micro*, Vol. 2 (1993b), pp. 289-344.

—. “Innovation and Market Value.” In R. Barrell, G. Mason, and M. O'Mahoney, eds., *Productivity, Innovation and Economic Performance*, Cambridge, England: Cambridge University Press, 2000.

— and Ziedonis, R.H. “The Patent Paradox Revisited: Determinants of Patenting in the U.S. Semiconductor Industry, 1980-94.” *Rand Journal of Economics*, Vol. 32 (2001), pp. 101-128.

—, Jaffe, A., and Trajtenberg, M. “Market Value and Patent Citations: A First Look.” Working Paper No. 7741, NBER, 2000.

—, —, and —. “The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools.” Working Paper No. 8498, NBER, 2001. Reprinted as Chapter 13 in Jaffe and Trajtenberg (2002)

Harhoff, D., Narin, F., Scherer, F.M., and Vopel, K. 1999. “Citation Frequency and the Value of Patented Inventions,” *The Review of Economics and Statistics* 81(3): 511-515.

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

— and Trajtenberg, M. *Patents, Citations and Innovations: A Window on the Knowledge Economy.* Cambridge: MIT Press, 2002.

—, —, and Fogarty, M. “The Meaning of Patent Citations: Report of the NBER/Case Western Reserve Survey of Patentees.” Working Paper No. 7631, NBER, 2000. Reprinted as Chapter 12 in Jaffe and Trajtenberg (2002).

—, —, and Henderson, R. “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations.” *Quarterly Journal of Economics*, Vol. 108 (1993), pp. 577-598.

Kortum, S. and Lerner, J. “Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?” *Carnegie-Rochester Conference Series on Public Policy*, Vol. 48 (1998), pp. 247-304.

Lanjouw, J.O. and Schankerman, M. “Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators.” *Economic Journal*, forthcoming.

Office of Technology Assessment and Forecast (OTAF). *U.S. Department of Commerce, Patent and Trademark Office, Sixth Report*. Washington, DC: US Government Printing Office, 1976.

Pakes, A. “On Patents, R&D, and the Stock Market Rate of Return.” *Journal of Political Economy*, Vol. 93 (1985), pp. 390-409.

—. “Patents as Options: Some Estimates of the Value of Holding European Patent Stocks.” *Econometrica*, Vol. 54 (1986), pp. 755-784.

— and Schankerman, M. “The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources.” In Z. Griliches, ed., *R&D, Patents, and Productivity*, Chicago: University of Chicago Press, 1984.

Quillen, C.D., Jr. and Webster, O.H. 2001. “Continuing Patent Applications and Performance of the U.S. Patent Office.” *Federal Circuit Bar Journal*, Vol. 11 (2001), pp. 1-21.

Scherer, F.M. “Firm Size, Market Structure, Opportunity, and the Output of Patented Innovations.” *American Economic Review*, Vol. 55 (1965), pp. 1097-1123.

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

Scotchmer, S. “Standing on the Shoulders of Giants: Cumulative Innovation and the Patent Law.” *Journal of Economic Perspectives*, Vol. 5 (1991), pp. 29-41.

Shane, H. “Patent Citations as an Indicator of the Value of Intangible Assets in the Semiconductor Industry.” Photocopy, The Wharton School, University of Pennsylvania, 1993.

Trajtenberg, M. “A Penny for Your Quotes: Patent Citations and the Value of Innovations.” *Rand Journal of Economics*, Vol. 21 (1990), pp. 172-187.

—, Henderson, R., and Jaffe, A. “University versus Corporate Patents: A Window on the Basicness of Invention.” *Economics of Innovation and New Technology*, Vol. 5 (1997), pp. 19-50.

Table 1

Sample Statistics for Patenting Firms, 1979-1988 - 12,118 observations

Variable	Mean	Median	Min.	Max.	Std. dev.	Share of within variance
Market value ($M)	916.33	77.41	.103	97,437	3,670	.10
Book value ($M)	914.21	65.16	.185	84,902	3,885	.03
Market-to-book value	1.73	1.09	.055	19.95	2.12	.34
R&D stock ($M)[a]	179.51	6.82	.002	22,130	885.6	.08
Patent stock	75.72	5.05	.033	5,085	291.6	.01
Citations stock[b]	569.82	38.34	.003	51,301	2,396	.02
D (R&D=0)	.16	.00	.0	1.00	.35	.05
R&D/Assets	.35	.16	.0	4.99	.57	.14
Patents/R&D[a]	1.25	.51	.0	252	5.32	.59
Citations/Patents	7.95	6.33	.0	222	7.22	.18
Past Citations stock	1.59	1.24	.0	33.38	1.70	.22
Past citations/total citations	.23	.20	.0	1.00	.17	.28
Predictable citations stock	7.95	7.17	2.25	64.76	3.44	.18
Residual citations stock	0	-1.09	-24.60	214.41	6.35	.28
Share of self-citations	.09	.06	.0	1.00	.13	.12

[a] For the 10,158 observations with R&D>0
[b] For 12,007 observations with Cites>0

Table 2
"Horse race" Regressions of R&D, Patents and Citations Stocks
Nonlinear model with dependent variable: log Tobin's q

	Sample: All Firms					
	Period: 1976-1984			**Period: 1985-1992**		
	11,711 observations, 3,100 firms			*15,605 observations, 3027 firms*		
R&D/Assets	1.736			.547		
	(.069)			*(.027)*		
Patents/Assets		.607			.710	
		(.042)			*(.049)*	
Cites/Assets			.103			.080
			(.006)			*(.004)*
D(R&D=0)	.029			.004		
	(.014)			*(.015)*		
D(PAT=0)		.120	.143		.144	.151
		(.013)	*(.013)*		*(.014)*	*(.014)*
R^2	.228	.110	.138	.116	.067	.087
Standard error	.699	.719	.707	.748	.769	.760
TotR&D/R&D	.635			.903		
PAT/R&D		.417			.306	
CITES/R&D*			2.784			2.686
Normalized coefficients**	***1.102***	***.253***	***.292***	***.494***	***.217***	***.210***

	Sample: Patenting Firms					
	1976-1984			**1985-1992**		
	10,509 observations, 1834 firms			*9,718 observations, 1843 firms*		
R&D/Assets	1.754			.563		
	(.082)			*(.033)*		
Patents/Assets		.599			.710	
		(.041)			*(.049)*	
Cites/Assets			.102			.080
			(.006)			*(.004)*
D(R&D=0)	.024					
	(.017)					
R^2	.231	.129	.180	.121	.103	.137
Standard error	.644	.685	.665	.729	.737	.722
TotR&D/R&D	.636			.903		
PAT/R&D		.458			.326	
CITES/R&D*			3.059			2.861
Normalized coefficients**	***1.116***	***.274***	***.321***	***.508***	***.231***	***.220***

Estimation method: nonlinear LS.
Heteroskedastic-consistent standard errors are shown in parentheses.
All equations include a complete set of year dummies.
*Ratio of Total R&D, Patents, or Citations to deflated R&D (in $1992M).
** Coefficients scaled by ratio of totals.

Table 3
Market Value as a Function of R&D, Patents, and Citations Stocks
Period: 1979-88; Sample: 1,982 Firms with patents>0; 12,118 observations
Nonlinear Model with dependent variable: log Tobin's q

	(1)	(2)	(3)	(4)	(5)
R&D/Assets	1.276	1.362	.926	1.261	1.135
	(.061)	*(.068)*	*(.047)*	*(.063)*	*(.053)*
Patents/R&D	.027	.030	.025	.028	.022
	(.006)	*(.007)*	*(.006)*	*(.007)*	*(.006)*
Cites/Patents		.052			
		(.004)			
Past Cites/Patents				-.015	
				(.007)	
Future Cites/ Patents				.060	
				(.004)	
Predicted Cites/ Patents					.012
					(.004)
Residual Cites/ Patents					.050
					(.003)
D(R&D=0)	.058	.057	.057	.087	.085
	(.019)	*(.019)*	*(.019)*	*(.019)*	*(.019)*
Dummies for # of cites per patent:*					
5 – 6 *(3,145 obs.)*			.006		
			(.018)		
7 – 10 *(3,993 obs.)*			.097		
			(.018)		
11 – 20 *(1,997 obs.)*			.353		
			(.023)		
> 20 *(573 obs.)*			.542		
			(.042)		
R^2	.222	.254	.255	.260	.260
Standard error	.685	.671	.670	.668	.668

* Base category: 0 – 4 cites per patent, 2,410 observations.
Estimation method: nonlinear LS.
Heteroskedastic-consistent standard errors are shown in parentheses.
All equations include a complete set of year dummies.

Table 4
Computing the Impact of the Knowledge Stocks on Market Value

Ratios	***Ratios evaluated at the:*** **Mean**	**Median**	**Ratio of totals**
R&D/Assets	.35	.16	.16
Patents/R&D	1.05	.35	.50
Cites/Patents	7.95	6.33	7.46
Semi-elasticities[a]			
$\frac{\partial \log Q}{\partial (R\&D/A)}$	.709 *(.026)*	.876 *(.037)*	.842 *(.036)*
$\frac{\partial \log Q}{\partial (PAT/R\&D)}$	.016 *(.004)*	.019 *(.005)*	.019 *(.004)*
$\frac{\partial \log Q}{\partial (CITES/PAT)}$	.027 *(.002)*	.033 *(.002)*	.032 *(.002)*

[a] Computed using the estimated coefficients in column (2) of Table 3. Heteroskedastic-consistent standard errors are shown in parentheses.

Table 5
Adding Industry Effects
Sample: 1,982 Patenting Firms, 1979-88 - 12,118 observations
Nonlinear Model with dependent variable: log Tobin's q

	(1)	(2)	(3)
D (Drugs)		.536 *(.028)*	.005 *(.102)*
D (Chemicals)		.026 *(.020)*	-.231 *(.061)*
D (Computers)		.311 *(.022)*	.361 *(.050)*
D (Electrical)		.166 *(.021)*	.093 *(.061)*
D (Metals & Machinery)		.015 *(.016)*	-.261 *(.047)*
R&D/Assets	1.362 *(.068)*	.686 *(.057)*	.883 *(.198)*
interacted with,			
Drugs			.561 *(.310)*
Chemicals			-.017 *(.333)*
Computers			-.575 *(.204)*
Electrical			-.343 *(.253)*
Metals & Machinery			.595 *(.241)*
Patents/R&D	.030 *(.007)*	.025 *(.006)*	-.020 *(.006)*
interacted with,			
Drugs			.120 *(.051)*
Chemicals			.059 *(.018)*
Computers			.078 *(.017)*
Electrical			.022 *(.006)*
Metals & Machinery			.070 *(.014)*
Citations/Patents	.052 *(.004)*	.036 *(.003)*	.014 *(.004)*
interacted with,			
Drugs			.065 *(.015)*
Chemicals			.048 *(.012)*
Computers			.014 *(.006)*
Electrical			.022 *(.011)*
Metals & Machinery			.037 *(.009)*
D (R&D=0)	.066 *(.019)*	.099 *(.018)*	.123 *(.020)*
R^2	.254	.292	.308
Standard Error	.671	.654	.647
Robust Wald Test for added effects (degrees of freedom)		503.5 *(5)*	142.6 *(15)*

Estimation method: nonlinear LS.
Heteroskedastic-consistent standard errors in parentheses.
All equations include year dummies.
The left-out category is miscellaneous (low-tech industries).

Table 6
Market Valuation of Self-citations
Sample: 1,982 Patenting Firms, 1979-88 - 12,118 observations
Nonlinear Model with dependent variable: log Tobin's q

	(1)	(2)	(3)	(4)
R&D/Assets	1.366	1.392	1.368	1.330
	(.076)	*(.079)*	*(.076)*	*(.074)*
Patents/R&D	.031	.031	.031	.029
	(.008)	*(.008)*	*(.008)*	*(.007)*
Cites/Patents	.051	.052	.043	.045
	(.004)	*(.004)*	*(.004)*	*(.004)*
Self-cites/Patents			.053	.067
			(.014)	*(.016)*
Self-cites/Total Cites		.218		
		(.093)		
[Self-cites/Patents] x log(patent portfolio)				-.026
				(.004)
D(R&D=0)	.066	.070	.073	.051
	(.019)	*(.019)*	*(.019)*	*(.019)*
R^2	.254	.254	.255	.258
Standard error	.671	.671	.670	.669

The log of patent portfolio has its mean removed, so the coefficient is relative to the average-sized firm. The range of log(portfolio) is 0 to 9.7, or 1 to 22,000.

Table A.1
Truncation Correction Weights

Year	Inverse weight
1964	2.119
1965	1.229
1966	1.063
1967	1.000
1968-89	1.000
1990	1.000
1991	1.003
1992	1.009
1993	1.034
1994	1.166
1995	2.230
1996	37.461

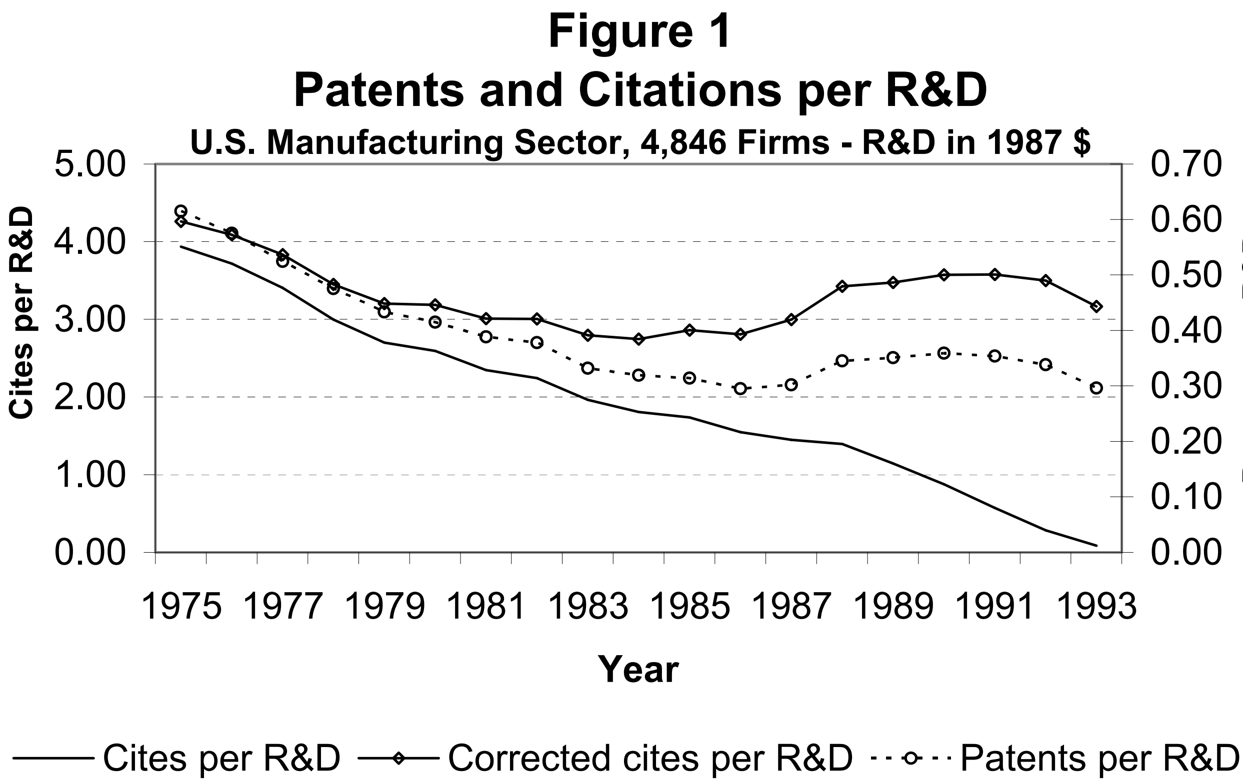


Figure 1
Patents and Citations per R&D
U.S. Manufacturing Sector, 4,846 Firms - R&D in 1987 $
5.00
4.00
3.00
2.00
1.00
0.00
Cites per R&D
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Patents per R&D
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
Year
Cites per R&D
Corrected cites per R&D
Patents per R&D

Figure 2
Splitting Citation Stocks into Predictable and Unpredictable Components

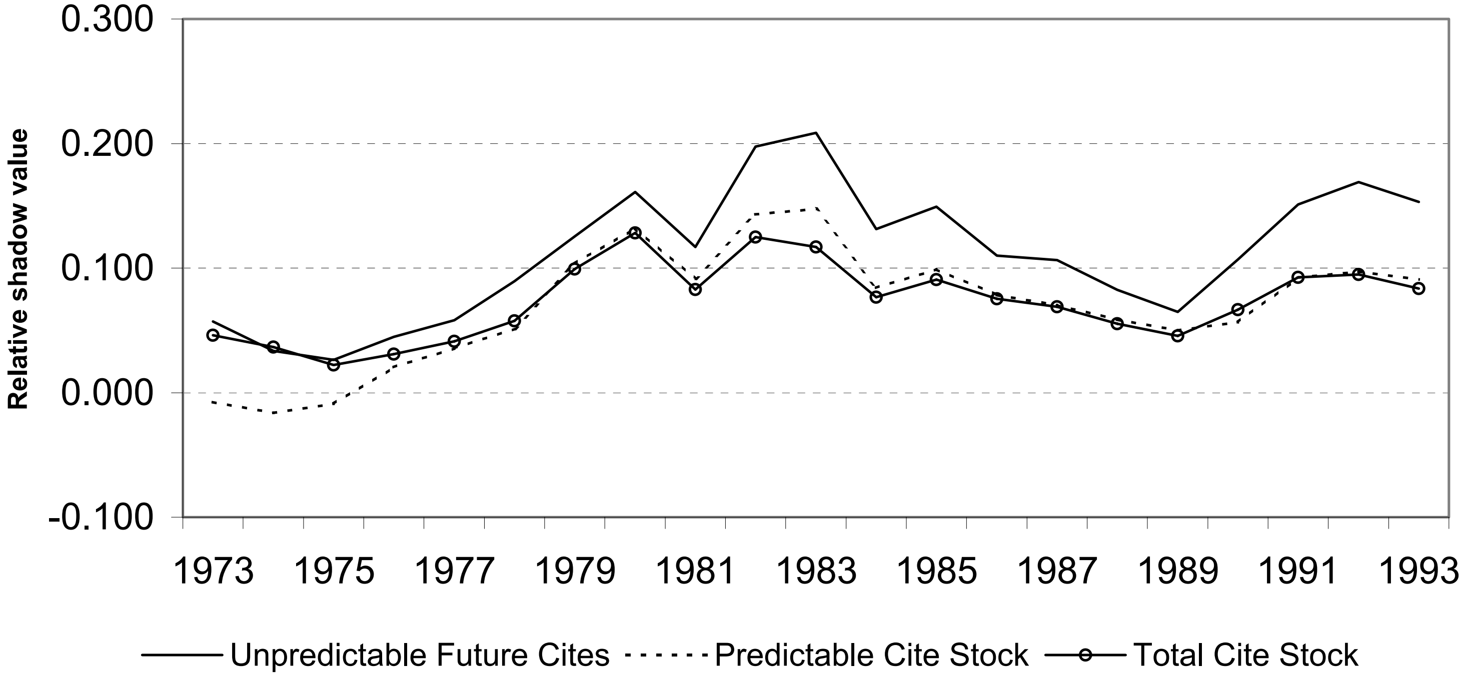


Figure A.1
Application-Grant Lag Distribution