EXHIBIT E

RAND Journal of Economics
Vol. 21, No. 1, Spring 1990

A penny for your quotes: patent citations and the value of innovations

Manuel Trajtenberg*

The use of patents in economic research has been seriously hindered by the fact that patents vary enormously in their importance or value, and hence, simple patent counts cannot be informative about innovative output. The purpose of this article is to put forward patent counts weighted by citations *as indicators of the value of innovations, thereby overcoming the limitations of simple counts. The empirical analysis of a particular innovation (Computed Tomography scanners) indeed shows a close association between citation-based patent indices and independent measures of the social value of innovations in that field. Moreover, the weighting scheme appears to be nonlinear (increasing) in the number of citations, implying that the informational content of citations* rises *at the margin. As in previous studies, simple patent counts are found to be highly correlated with contemporaneous R&D; however, here the association is within a field over time rather than cross-sectional.*

1. Introduction

■ The study of technological change has been hampered all along by the scarcity of appropriate data and, in particular, by the lack of good indicators of innovation having a wide coverage. Patents would seem to be the one important exception, since they are the only manifestation of inventive activity covering virtually every field of innovation in most developed countries and over long periods of time. Yet, their use in economic research has not lived up to expectations primarily because patents exhibit an enormous variance in their "importance" or "value," and hence, simple patent counts cannot be very informative of innovative "output" (which is usually what we are after).

The goal here is to readdress this problem by examining the usefulness of patent indicators in the context of a particular innovation, Computed Tomography scanners, one of the most important advances in medical technology of recent times. The central hypothesis is that patent citations (i.e., references to patents appearing in the patent documents themselves), long presumed to be indicative of something like technological importance, may be informative of the economic value of innovations as well. Indeed, patent counts weighted by a citations-based index are found to be highly correlated (over time) with independent measures of the social gains from innovations in Computed Tomography. On the other hand, as in the previous literature, simple patent counts are found to be indicative only of the input side, as reflected in R&D outlays. Beyond establishing their role as indicators, the

* Tel-Aviv University and Harvard Business School.
I am grateful to Tim Bresnahan, Zvi Griliches, an anonymous referee, and to participants in the productivity workshop of the NBER Summer Institute for helpful comments on an earlier draft.

findings suggest that citations may reflect a sort of causal relationship between citing and cited patents, which is consistent with the view of innovation as a continuous and incremental process, punctuated by occasional breakthroughs.

The article is organized as follows. Section 2 lays out the basis for the use of patent citations. Section 3 offers a first view of the patent data in Computed Tomography, and Section 4 briefly describes the measures of the value of innovations in this field borrowed from a previous study (Trajtenberg, 1989). The tests of the main hypotheses are conducted in Section 5, and Section 6 offers some closing thoughts. Finally, in view of the fact that the number of citations per patent decreases drastically over time, I present in the Appendix a statistical procedure to test for "age versus importance" and for truncation effects.

2. Using patent counts and patent citations

■ It has long been thought that the detailed information contained in the patent documents may have a bearing on the importance of the innovations disclosed in them and that it may therefore be possible to construct patent indicators that could serve as proxies for the value of innovations.[1] Up to now, though, virtually the only patent measures used in economic research have been simple patent counts (henceforth SPC), that is, the number of patents assigned over a certain period of time to firms, industries, countries, *etc.*

The body of evidence that has accumulated since Schmookler (1966) indicates fairly clearly that SPC are closely associated with the input side of the innovative process, primarily with contemporaneous R&D expenditures in the cross-sectional dimension (Griliches, 1984). On the other hand, the few attempts to relate those counts to value indicators (e.g., the market value of innovating firms) have been largely unsuccessful. (See, for example, Griliches *et al.,* 1988.) As suggested above, those findings are hardly surprising, considering that patents vary enormously in their technological and economic significance. Thus, the mere counting of patents at any level of aggregation cannot possibly render good value indicators: simple patent counts assign a value of one to all patents by construction, whereas their true values exhibit a very large variance. Furthermore, there is substantial evidence to the effect that the distribution of patent values is highly skewed toward the low end, with a long and thin tail into the high-value side. As Scherer (1965) notes, those Pareto-like distributions might not have finite moments and, in particular, they might not have a finite variance; clearly, that would make the use of patent counts as proxies even more problematic. It is important to emphasize that those problems are inherent to the patent system as such,[2] and therefore definite solutions can hardly be expected.

An idea that has often been suggested is to use patent *citations* as an index of the importance or value of patents,[3] i.e., to count the number of times that each patent has been cited in subsequent patents and use the number to compute weighted patent counts. (I am referring to the citations appearing on the front page of patents under *References Cited.*) The potential significance of patent citations can be inferred from the following quotation:

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 might anticipate the claimed invention and

[1] By value I mean the social benefits generated by the innovation in the form of the additional consumer surplus and the profits stemming from the innovation. The "value," "output," and "magnitude" of innovations are taken to mean exactly the same thing. (For a detailed discussion of those concepts, see Trajtenberg (1990).)

[2] This is so because the importance of a patent—however defined—can hardly be assessed *ex ante* and because it is not the task of patent examiners to make sure that the patents granted are of comparable worth.

[3] This idea draws from the extensive use of citations from and to scientific publications in the context of bibliometric studies. (See, for example, Price (1963).)

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)

Moreover, there is a legal dimension to patent citations, since they represent a limitation on the scope of the property rights established by a patent's claims, which carry weight in court. Equally important, the process of arriving at the final list of references, which involves the applicant and his attorney as well as the examiner, apparently does generate the right incentives to have all relevant patents cited, and only those. (See Campbell and Nieves (1979).) The presumption that citation counts are potentially informative of something like the technological importance of patents is thus well grounded.

The question is whether citations counts may also be indicative of the (*ex post*) value of the innovations disclosed in the cited patents.[4] This can only be answered empirically, but one can advance some further arguments that would strengthen the prior. Most patents cited are referenced in patents issued within the same narrowly defined field of innovation as the cited patents ("within citation"). The very existence of those later patents attests to the fact that the cited patents opened the way to a technologically successful line of innovation. Moreover, it presumably attests also to economic success (at least in expected value terms), since those subsequent patents are the result of costly innovational efforts undertaken mostly by profit-seeking agents. Given that citations to a patent are counted for a period of a few years following its issuance, there should be enough time for the uncertainty regarding the economic value of the innovation to resolve itself. Thus, if citations keep coming, it must be that the innovation originating in the cited patent had indeed proven to be valuable.

Whatever their merits, patent citations have rarely been used in economic research[5] probably because it used to be quite difficult to obtain the necessary data (i.e., the frequency of citations for each patent studied). Today, however, this is easily done with the aid of computerized search techniques, such as those offered by DIALOG or BRS. More important, the significance of a citations-based index can be ascertained only by relating it to independent measures of the value of innovations; given the scarcity of such measures, no firm conclusions could have possibly been drawn from citations-based statistical findings.[6] It is here that the advantage of having estimates of the social gains from innovation in Computed Tomography scanners from a previous study proved to be crucial. A further advantage is that both the patent counts and the value measures used to validate them refer in this case precisely to the same stretch of innovative activity, i.e., to advances in a carefully circumscribed product class and time period. Thus, the usual problems that arise when trying to match information belonging to disparate units (as often happens in this context) are altogether absent here.

Granted the use of citations, the next issue is how to go about constructing a sensible weighting scheme. A straightforward possibility is to weight each patent i by the actual number of citations that it subsequently received, denoted by C_i. Thus, if I were to compute

---

[4] This clearly need not be the same as technological importance: the latter could be thought of as having to do only with the supply side of innovations, whereas value obviously reflects a market equilibrium.

[5] An exception is Lieberman (1987). Other studies using patent citations in a related way (i.e., seeking to approximate something like importance) but not quite in the realm of economics proper include Carpenter *et al.* (1981), Ellis *et al.* (1978), and Narin and Wolf (1983).

[6] A series of articles in the press proclaimed some time ago that ". . . Japanese Outpace Americans in Innovation," on the basis of findings showing that Japanese patents have been cited more often than U.S. patents. (See *New York Times*, (March 7, 1988) and *Time*, (March 21, 1988); the study cited was conducted by Computer Horizons, Inc. for the National Science Foundation.) While that may be so, the mere finding of a higher frequency of citations does not and cannot prove anything by itself.

an index of weighted patent counts (WPC) for, say, a given product class in a given year, t, I would have,

$$WPC_t = \sum_{i=1}^{n_t} (1 + C_i), \qquad (1)$$

where n_t is the number of patents issued during year t in that product class. This linear weighting scheme then assigns a value of one to all citations and all patents. Lacking more information on the citation process, this is a natural starting point but certainly not the only possibility; in fact, in Section 5, I compute a nonlinear index that allows for the existence of "returns to scale" in the informational content of citations.

3. Patents in Computed Tomography: a first look

■ Computed Tomography (CT) is a sophisticated diagnostic technology that produces cross-sectional images of the interior of the body, allowing the visualization of a wide range of organs with great accuracy. It has been hailed as one of the most remarkable medical innovations of recent times, comparable to the invention of radiography. Originally developed at the British firm EMI in the early seventies, CT soon attracted some twenty other firms worldwide, and the fierce competition that ensued brought about a breathtaking pace of technical advance. The diffusion of the new systems also proceeded very quickly: first introduced in the U.S. in 1973, by 1985 almost 60% of hospitals (community hospitals with more than 100 beds) had at least one system installed. The pace of innovation subsided in the late seventies as the technology matured and ceded its dominant place to new developments, particularly to Magnetic Resonance Imaging.

Searching in the PATDATA database (available through BRS), I located and retrieved all U.S. patents granted in Computed Tomography from the very start of the field in 1971[7] up to the end of 1986, totalling 456 patents.[8] I am quite certain that this is indeed the complete set and that it includes patents in CT only.[9] Clearly, having a clean set of patents, and hence clean patent-based indicators, is crucial in order to assess the usefulness of those indicators; otherwise, it would be impossible to tell whether the results are spurious (due to errors of measurement) or reflect real phenomena.

As is by now standard practice, patents are dated according to their application, rather than granting, date. Examining the distribution of lags between these two dates, I concluded that the patent data comprise virtually all patents applied for up to the end of 1982, 96% of the patents applied for in 1983, and smaller percentages of those applied for in later years. Thus, the analysis will be confined to the period of 1972–1982, although the citations appearing in the 1983–1986 patents will be taken into account as well.

Citation counts can be done in two ways: counting all citations or just those appearing in the set of patents belonging to the same field. In the "within referencing" case, the citation counts will be associated with the value of the innovations in the specific technological field to which they belong. On the other hand, an all-inclusive count will presumably capture

---

[7] In this case, it was easy to identify the very first patent: the origin of Computed Tomography is unequivocally associated with its invention by G. Hounsfield, as described in his U.S. patent #3778614, applied for in December 1971. Since there were no patents in CT in 1972, I shall treat this first patent as if it had been applied for in January 1972, rather than in December 1971, in order to avoid an unnecessary discontinuity in the data points.

[8] See Trajtenberg (1990) for a detailed discussion of the long-standing classification problem (i.e., how to match patents to economic categories) and for the advantages of using computerized search techniques in order to tackle it.

[9] The computerized search actually produced 501 patents, but 45 of them were eliminated after a careful examination of their abstracts; thus, I am certain that all the patents included do belong to CT. Furthermore, I am confident that the set includes all the relevant patents, since I was able to cross-check with other sources, including listings of patents from the manufacturers of CT scanners.

the value "spilled-over" to other areas as well. Given that the measures of innovation to be used in conjunction with the patent data refer to the gains from advances in CT as such, with no attempt to account for spillovers, I shall use the within referencing count.[10]

The first two columns of Table 1, graphically displayed in Figure 1, show the basic patent data to be used throughout. Note the smooth, cycle-like path followed by the yearly count of patents: it rises quite fast after 1973, peaks in 1977, and then declines steadily, carrying forward a thin tail. Notice also that the weighting scheme strongly influences the shape of the distribution, shifting it back toward the earlier period. In fact, the difference in the means of the two distributions is seventeen months; that is, the weighting scheme centers the action in mid-1976 rather than in late 1977. Given the very fast pace at which the CT technology evolved and that the period is just eleven years long, a difference of one and one-half years in the means is certainly very significant. Clearly, this shift is due to the fact that earlier patents were cited more frequently than later ones: as Table 1 shows, the average number of citations per patent went down from 72 to less than one, and the percentage of patents with no citations increased from zero to 92%.

The question is whether the observed citation frequencies are to be regarded as real phenomema, presumably reflecting something like the importance of patents, or just as statistical artifacts, induced by the mere passage of time. Two concerns arise in this context. First, it could be that older patents are cited more often simply because they have had more opportunities to be cited, since they precede a larger set of patents that could cite them. Second, given that CT is an ongoing technology (albeit already mature), it is quite certain that additional patents have been issued since the time of the search and that more will be granted in the future. Thus, the data set is necessarily truncated, which might bias downward the citation counts of recent patents.

TABLE 1 **Patents in Computed Tomography: Counts and Citations by Year**

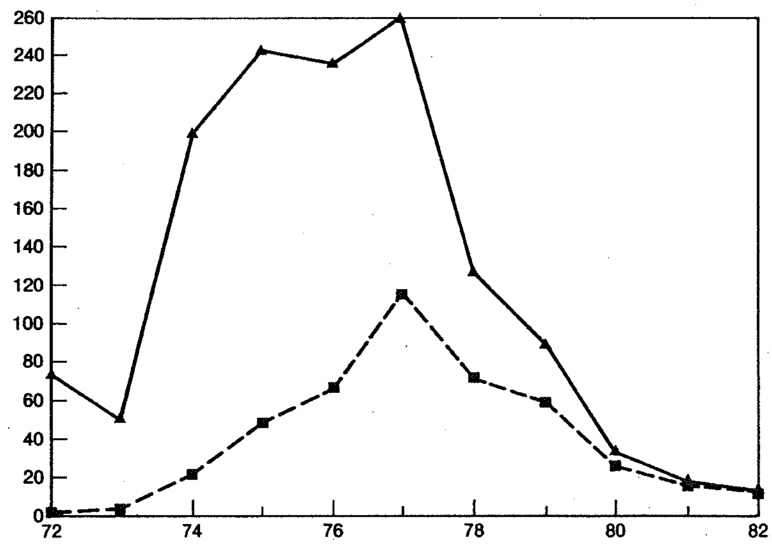
	Patents		Citations		
				Percentage of Patents with:	
Year	Simple Counts (*SPC*)	Weighted by Citations (*WPC*)	Average No. per Patent	0	5+
1972	1	73	72.0	0.0	100.0
1973	3	50	15.7	0.0	100.0
1974	21	199	8.5	4.8	76.2
1975	48	242	4.0	12.5	47.9
1976	66	235	2.6	21.2	22.7
1977	115	260	1.3	45.2	11.3
1978	71	126	0.8	54.9	4.2
1979	59	88	0.5	66.1	0.0
1980	26	33[a]	0.3	84.6	0.0
1981	15	18[a]	0.2	86.7	0.0
1982	12	13[a]	0.1	91.7	0.0
1983[b]	13	14	•	•	•
1984[b]	6	6	•	•	•
All	456	1357	2.1[c]	45.1[c]	16.2[c]

[a] These figures are slightly biased downwards. (See the Appendix).
[b] These are partial figures.
[c] These are averages.

[10] In this case, it would not have mattered much which count was used: in a sample of 30 patents in CT, the correlation between the two counts was found to be .99. Similarly, Campbell and Nieves (1979) report a correlation of .73 between what they called "in-set" and total citations for some 800 patents in the field of catalytic converters.

FIGURE 1

PATENTS IN CT: SIMPLE COUNTS, AND COUNTS WEIGHTED BY CITATIONS



■ Simple patent count (Source: Table 1, Column 1).

▲ Count weighted by citations (Source: Table 1, Column 2).

These are serious *a priori* objections that may arise whenever one tries to attach any meaning to citations data, and therefore deserve careful scrutiny.[11] The Appendix analyzes them in detail and shows that neither age nor truncation could possibly account for the observed distribution of citation counts. The issue of age is tackled by constructing a hypothetical "iso-important" distribution of citations and testing it against the observed distribution with the aid of a χ^2-test: the null hypothesis that older patents received more citations just because of the passage of time is rejected by a wide margin. As for the effect of truncation, the magnitude of the biases is estimated by extrapolating from the observed distribution of citation lags and application-granting lags. The main finding is that a bias does exist, but the absolute expected number of missing citations to recent patents is very small and hence could not possibly affect the results of the statistical analysis based on citation counts.

4. Estimates of the value of innovations in Computed Tomography

■ As stated earlier, I intend to relate patent counts weighted by citations to independent measures of innovation in CT taken from an earlier study (Trajtenberg, 1989). The basic idea behind those measures is as follows. Consider a technologically dynamic product class as it evolves over time, and assume that the different brands in it can be described in terms of a small number of attributes. Product innovation can then be thought of in terms of changes in the set of available products, both in that new brands appear and that the qualities of existing products improve. Applying discrete choice models to data on sales per brand

[11] The issue of age versus importance (closely related to Price's "immediacy factor") has commanded a great deal of attention in the scientometric literature. However, to the best of my knowledge, it has not yet been addressed with rigorous statistical tests. (See, for example, Line (1970) and Campbell and Nieves (1979).)

and on their attributes and prices, one can estimate the parameters of the demand functions and, under some restrictions, of the underlying utility function. The social value of the innovations occurring between two periods can then be calculated as the benefits of having the latest choice set rather than the previous one in terms of the ensuing increments in consumer and producer surplus. That is, given an estimated social surplus function, $W(\cdot)$, and the sets of products offered in two successive periods, S_t and S_{t-1}, the value of innovation would be measured by $\Delta W_t = W(S_t) - W(S_{t-1})$.

The model used to estimate the function $W(\cdot)$ was the multinomial logit model, rendering the well-known choice probabilities $\pi_j = \exp\phi(z_{jt}, p_{jt})/\sum_i^m \exp[\phi(z_{it}, p_{it})]$, where z is the vector of attributes, p is price, m is the number of alternatives in the choice set, and $\phi(\cdot)$ is the branch of the indirect utility function related to the product class in question. Integrating those probabilistic demand functions, one obtains measures of consumer surplus of the form[12]

$$W_t = \ln\{\sum_i^{m_t} \exp[\phi(z_{it}, p_{it})]\}/\lambda_t, \tag{2}$$

where λ stands for the marginal utility of income. The differences ΔW_t are then computed from (2) for every pair of adjacent years. Noting that ΔW_t refers to the gains accruing to the representative consumer, I also computed the total gains associated with the innovations at t, using

$$TW_t = \Delta W_t\left[n_t + K(\sum_{\tau=0}^{t} \Delta W_\tau)\int_{t+1}^{\infty} f(\tau)e^{-r(\tau-t-1)}d\tau\right] \equiv \Delta W_t(n_t + n_f), \tag{3}$$

where n_t is the number of consumers at t, $K(\cdot)$ is the ceiling of the diffusion curve (that shifts up as a consequence of successive innovations), $f(\cdot)$ is the diffusion path, and r is the interest rate. Thus, TW multiplies ΔW_t by the current and future number of consumers that benefit from the innovations at t, the latter being assessed on the basis of the observed diffusion process.

For the earlier study, I gathered a comprehensive data set on CT scanners, including the prices and attributes of all scanners marketed in the U.S. since the inception of CT in 1973 up to 1982, and details of all sales to U.S. hospitals (over 2,000 observations). Applying the methodology just sketched to these data, I obtained yearly estimates of ΔW_t and of TW_t, as shown in Table 2 along with other data on CT. Note that ΔW_t and TW_t are very large at first and then decline sharply, carrying forward a thin tail. Thus, the bulk of gains from innovation were generated during the earlier years of the technology even though the action in terms of R&D and entry peaks later on.

5. Patents as indicators of innovation: the statistical evidence

■ The question to be addressed here is, then, To what extent can patent-based indices (denoted by P) serve as indicators of the value of innovations as measured by ΔW (or

---

[12] In the present case, ΔW was confined to changes in consumer surplus, since the net aggregate profits of CT manufacturers were actually nil over the period studied. This may raise questions as to the rationale of the expected link between ΔW and patent indices, since, from the point of view of the incentives to innovate (and hence, to patent), what counts are the benefits appropriated by the firms, not consumer surplus. This is true *ex ante*; that is, firms innovate because they expect to be able to appropriate a great deal of the social surplus, but competition may significantly reduce (*ex post*) that portion. (Recall that competition in the CT market was indeed very intense during the period studied.) In that case one would still expect to find a significant correlation between total surplus and, say, *WPC*, even if the former is made just of consumer surplus.