European Management Review (2008) 5, 85–89

palgrave-journals.com/emr

Comment

The value of European patents

Michelle Gittelman

Rutgers Business School, Newark-New Brunswick, New Jersey, USA

Correspondence: Michelle Gittelman, Rutgers Business School, Newark-New Brunswick, NJ, USA.
E-mail: michelle.gittelman@business.rutgers.edu

Abstract

This paper is an important contribution to the growing literature on the meaning of patent data. Its core findings should become part of the vernacular in empirical research using patents as measures of innovation. A startling finding is that commonly used patent-based proxies for value capture less than 5% of variation in value as reported by inventors. The results stand as a call for further exploration of the complex social and institutional processes by which patents are made. In this comment, I reflect on the main findings and their implications for what we know – and have yet to learn – about the value of patents and the meaning of patent data in innovation studies.

European Management Review (2008) **5,** 85–89. doi:10.1057/emr.2008.12
Keywords: patents; patent claims; patent citations; value of patents; innovation

'The Value of European Patents' is an important milestone in a growing body of work on the meaning of patent data. The authors – Alfonso Gambardella, Dietmar Harhoff, and Bart Verspagen (GHV) – are three economists who have each made major contributions to the literature on technological innovation. This article is a careful analysis of data collected in the PAT-VAL survey, a large-scale, multi-country project in which inventors were asked to report on the monetary value of their patents. Survey data of this sort are very difficult to collect, particularly on such a large scale, and the few survey-based studies we have are like bits of amber on a beach, rare and precious. The findings of this article challenge us to question some common assumptions about the value of patents and their meaning as data, and the core results should become part of the vernacular of patent-based research.

In addition to closing some gaps in our understanding of patent values and patent data, Gambardella *et al*. also open up a space that may be shared by economists using large-scale data sets and sociologists who study the micro-level processes and socio-technical networks through which inventions and patents are made. The statistical results in this article tell us that patent data can explain only a very small portion of value as reported by inventors. This large gap invites us to develop a greater understanding of how patents are made, as they are complex documents constructed by heterogeneous groups of actors, each reflecting different skills, strategies, and value calculations.

An analogy here can be made between empirical researchers in technological innovation and historians of the Victorian era, who mine documents for detailed demographic information that was beginning to be collected and encoded at that time. Certificates of birth, marriage and death, police and hospital records generated data on social and economic phenomena that are not available for earlier eras, when codification was much spottier. Similarly, patent documents encode remarkably fine-grained information about inventions: written claims and abstracts; uniform codes for technology classes; the names of individual inventors and firms, along with their precise geographic locations; the identities of lawyers and patent examiners; and backward and forward citations. Tracing an invention's technological antecedents and impact on future inventions, each patent citation contains the same abundance of bibliometric minutiae, generating an endless regress of dyadic linkages that ends only when the data are no longer recorded. The ability to electronically retrieve patent documents and extract their components on a very large scale has opened up a rich vein of empirical research in innovation studies, and the number of papers using patent data has blossomed in the past two decades. Arguably our most valuable recent insights on the diffusion of knowledge across time, organizational boundaries, and space have been furnished by studies that utilize patent data.

Yet like historians, we are faced with the problem that encoded documents can only proxy – with substantial error – for the messy reality that they codify. Much is left unrecorded, and what is recorded is often recorded poorly. As users of these data, we would like to know the extent to which our measures actually mean something, if only to keep us honest and direct our search where the answers lie, rather than where the light of the data shines most brightly. It is one thing to observe that inventors patented for multiple organizations, but does that mean that their conversations conveyed meaningful information across those organizations? Does a citation mean that learning has occurred, or was it added by a patent examiner to deny an applicants' claim to novelty? Our interpretations of

results rest squarely upon the answers to such questions, and so it is important that we produce research that links our readily available patent data with real information that is much more arduous to collect.

Herein lies the value of the present study. It is an important addition to our too-small library of published studies that try to uncover the usefulness of patent data as measures of innovation.[1] It makes three major contributions: (1) it estimates from surveys of inventors the 'all-in' monetary value of individual patents, (2) it specifies the distributional properties of these values across a large sample of patents, and (3) it estimates statistically the degree to which measures that are commonly employed in the empirical literature (backward and forward citations, claims, renewals, and international scope) correlate with estimated private economic value. I discuss the first two points jointly in the next section and conclude with a section of remarks on the third.

Patent values and the value distribution

The paper first presents the distribution of patent values as reported by inventors, and carefully validates those estimates. At first glance, it would seem that point estimates of the monetary values of patents would be of greater interest to policy makers than innovation researchers, since many research questions can be sufficiently answered with ordinal rather than their cardinal value estimates. But the reported values and their distribution raise questions about the assumption that patents are good proxies for patented and unpatented ideas.

Common wisdom holds that, while a few inventions and patents are very valuable, the majority are worthless. Indeed, the serendipitous nature of the invention process would lead us to expect that most ideas will have little private value, and that a few will be 'hits,' and prior research establishes that patents exhibit a highly skewed value distribution; about half of all patents receive less than two citations, less than 2% are ever litigated; about one half to one third lapse without being renewed (e.g., Hall *et al*., 2005; Lemley and Shapiro, 2005).

With regard to monetary returns Gambardella *et al*. find, as expected, a long right tail of very valuable inventions, but what is unexpected is that the distribution as a whole is shifted further to the right than our priors would lead us to expect: the median patent in the sample is worth 650,000 euros, and the left tail is centered over levels that are surprisingly high (somewhere between 30,000 and 100,000 euros). The left tail indicates that even the least valuable patents carry respectable economic weight, more than might be expected if randomly drawn from a pool of patented and unpatented inventive ideas.

The data are consistent with the idea that applicants know *ex ante* which inventions are likely to be valuable and select inventions for patenting accordingly. From a measurement perspective this selection into patenting is a problem, since we do not observe those characteristics that caused inventions to be selected into patenting in the first place. It is difficult to claim that we study the sources of valuable inventions when value and/or novelty is the minimum threshold for inclusion into our data sets. And yet the fact that applicants know something about their inventions before they choose to patent is a simple behavioral assumption, since most inventions are not made in isolated laboratories but are developed by teams of people working in firms who think hard about the marketability, technical viability, and competitive landscape of a given invention. The model of a team of dogged researchers churning out ideas with a random distribution of characteristics, occasionally hitting upon the right combination to produce the odd success, is far from the reality of firms that carefully manage the R&D process with a fair sense about which ideas will hold their own as a technology, in a market, or in a courtroom.

Another way selection may operate is if applicants file patents randomly but ideas that do not meet the threshold of novelty and non-obviousness are denied a patent, satisfying our expectation that the patent office screens applications to ensure that patents represent truly inventive ideas. My coauthors, Juan Alcacer and Bhaven Sampat, and I have found evidence of this behavior, in that firms that receive the highest number of patents are more likely to be granted patents with no applicant prior art, an indication of low effort in preparing patents on the part of applicants (Alcacer *et al*., 2008). These applicants may use the patent office to outsource the quality screening process, rather than invest in screening in-house. Schneider (2007) finds that large applicants wait for a signal from the patent office before withdrawing their applications, another indication of idea screening by the patent office. The relatively high values on the left side of the distribution speaks to the operation of unobserved selection, by applicants or the patent office, that will confound causal explanations of the determinants of valuable ideas.

A second implication of the patent value distribution reported by Gambardella *et al*. is that patents are very valuable as strategic tools, a topic that has been the focus of much recent research (e.g., Arora *et al*., 2008). It is an innovation of this paper (and the PAT-VAL survey more broadly) that it reports patent monetary value as the all-in present value of an option that incorporates expected profits from a complex range of possible uses: exploitation, sale, blocking, rent extraction from owners of related patents, and rents foregone from selling to others who might subsequently block. The values reported here are thus a composite derived from use, trade, and leverage, from which we may gain many important insights.

Again, the left tail of the value distribution carries intriguing possibilities for future investigation. At the far right tail – where we find products worth hundreds of millions of euros – are many pharmaceutical products, where practically all value is likely to be captured in the patent premium (the incremental value of the monopoly conferred by the patent). We may hypothesize that as we travel from right to left on the value distribution the importance of monopoly rents declines as a relative share of value and the proportion afforded by strategic uses increases: blocking, patent pooling, cross-licensing, patent opposition, or litigation, etc. These issues are particularly salient in technologies that rely on assembly of bundles of inventions distributed among multiple patent owners. The fat left tail suggests that in the aggregate the value of these strategic uses is considerable.

While we still cannot quantify these effects, the shape of the distribution speaks to the fact that, like knowledge itself, a patent is a heterogeneous resource, and we should speak of patent values, rather than a unitary value contained in a patent. The data reported here allow us an opportunity to begin to unpack this heterogeneity. The work of Ashish Arora and his colleagues[2] has highlighted that strong patents are valuable because they allow for active trade in technology through licensing. Licensing markets in turn encourage the vertical dis-aggregation of industry value chains, as small firms with strong patents can specialize in research and out-license their inventions to large firms with complementary assets in marketing and distribution. The degree to which technology markets emerge and substitute for internal firm hierarchies depends on a complex mix of institutional, organizational, and technological factors, including local conditions for knowledge-based entrepreneurship. In this regard, these data can yield rich comparative studies, and it would be very instructive to collect similar data for the United States, where there is a great deal of institutional liquidity in licensing markets, and many possibilities for organizing firms as vehicles for exploiting patented knowledge to compare to the European data collected here.

Patent citations and patent value

The heterogeneous nature of patent value is further underscored in the next major set of results, which are the estimated conditional correlations between monetary values on the one hand and patent-based measures of value on the other. Here, there is good news and bad news for users of patent data in innovation studies. The good news is that bibliometric measures commonly used to indicate patent value – backward and forward citations, claims, renewals, and number of countries in which the patent is issued – do correlate with monetary values, and the correlations are highly significant. The bad news is that the magnitude of the explained variance is extremely low. Together, they jointly explain only about 3% of the variance in reported value. This is especially troubling for those measures that are in the control of patent applicant (backward references, renewals, and number of countries). Presumably, these should reflect the *ex ante* expectations of applicants about monetary value, and we do not have an explanation for why applicant effort/cost should explain such a small proportion of estimated monetary value. As the authors point out, the bibliometric measures correlate with private value, but with significant noise. This is a modest understatement, and the low explained variance should become a standard reference point in empirical research using bibliometric indicators to measure private value.

The finding that forward citations explain so little of the variance in private monetary value is puzzling. Citations represent diffusion, and diffusion is a form of value, though not necessarily private value. Forward citations have held up much better than backward citations in tests of their legitimacy as indicators of patent impact and knowledge flows, respectively. However, we discover here that not only do they explain very little of the variance in estimated value, there is noise at both tails of the distribution: about one-quarter of patents with no citations are ranked at or above the (roughly) top 20 percentage of patents in terms of their value, while two thirds of the most highly cited patents are ranked below that mark![3]

This is disappointing news to those (myself included) who find it convenient to use forward citations as a measure of the economic value of an invention. But perhaps the news is not so bad, for in fact citations are generated by multiple sources that have different motives to cite. A firm may cite a rivals' patent to show non-obviousness of its own invention; a patent examiner may define valuable citations as patents that afford economies of scale in search time across a large number of applications. A firms' attorneys may decide to publish a patent that the firm never intends to exploit except to create prior art that denies a rival claim to novelty or non-obviousness. Learning, blocking, convenience, and ceremonial acknowledgement reflect different motives to cite, and these may not have high correlations with private monetary value as reported by inventors.

This raises the question of what we measure when we pool citations. In work with Juan Alcacer (Alcacer and Gittelman, 2006) we disaggregate backwards citations, which are commonly used to measure knowledge flows between inventors. The standard assumption in the literature is that citations added by patent examiners – which presumably do not measure knowledge flows between inventors – add random noise to applicant citations, such that pooled citations (by examiners and applicants) may be considered a good, though noisy, signal of knowledge flows. Our results show that this assumption is incorrect. In fact, examiner citations and applicant citations track each other so closely along several dimensions commonly used to measure knowledge flows that examiner citations are hardly random with respect to applicant citations. This means that *statistically* inferences from pooled citations are not biased – but the close tracking raises conceptual questions about the validity of applicant citations as measures of knowledge flows. The fact that examiner citations are more likely than applicants to add self-citations than drives home this point.

With regard to highly cited patents, unobserved heterogeneity in terms of why the patent is being cited and by whom will likely affect the relationship between citations and patent value, and unpacking this heterogeneity seems a fruitful way to begin to understand the GHV results. We know that applicants and examiners have different 'tastes' for citing patents: in unpublished work with Juan Alcacer (Alcacer and Gittelman, 2004), we find that the most highly cited patents earn their status by being cited by examiners or applicants, but not both: examiners and applicants select different sets of patents for frequent citation. We also know that different types of citations have different meaning for value: Hall *et al.* (2005) find that self citations are more valuable to firm market value than citations by others, and Hegde and Sampat (2005) find that citations by examiners are more highly correlated with patent renewals than forward citations by other patent applicants.

It is tempting to speculate about the specific behaviors that could generate GHV's intriguing result that forward citations explain so little private value. Since their sample is composed of European patents, it is likely that the majority

of forward citations are made by patent examiners, since in Europe all citations are added by patent examiners. This could drive their result in several ways. Examiners – technical specialists in their field – frequently choose citations that are broad and general enough to teach a large swath of the 'art,' or current state of knowledge in a field. It is a time-efficient strategy in a world that is awash in prior art, but such broad patents may not yield much in terms of of private monetary value.

Over and above any technical characteristics, idiosyncratic variation across examiners may also be important to the GHV results. The job of the examiner is to assess the patentability of an invention in light of what is already known (the prior art). Examiners are not computers, and we may expect large idiosyncracies in their search routines and citation choices. In their study of US patent examiners, Cockburn *et al.* (2004) find that the individual examiner effects accounted for surprisingly large variation in forward citations received by patents, even after controlling for technology class, age, and other characteristics that could influence citations. Moreover, they find that examiners whose patents attract many citations are also more likely to have their patents subsequently found invalid in the courts. In part, this reflects the fact that US patent examiners are civil servants whose careers are not influenced by the economic or legal value of the patents they approve. It is also consistent with the GHV result, in showing that examiners frequently approve patents that are convenient to cite but worth relatively little as intellectual property.

Finally, applicants themselves could play a role in the GHV finding of a large disconnect between forward citations and value. Diffusion is a double-edged sword, and firms may reap publicity benefits but little else from widely-cited patents. Patent attorneys, who understand well features that make a patent easy to cite, could calibrate their efforts accordingly, designing patents on inconsequential inventions to be highly cited – a form of gift-giving to examiners and the technological communities in which their clients operate – while writing patents with less-citable features for more important inventions. There is value in obscurity, and the legal skills to achieve it are a marketable service. An example of such a strategy is to apply for multiple patents on a single invention, so that no single patent attracts many citations, and the invention as a whole is difficult for others to re-assemble from the pieces. These are but speculations for the GHV results; we are left with the need for more research about the micro-level processes by which patents are written and citations are made.

Again, we are back to the problem of the historian seeking to reveal complex phenomena from cross-sectional demographic data that are recorded at pre-specified times: birth, death, arrest. If we lift the rock on the work that goes into writing and citing a patent, we see a world of busy life wriggling beneath: scientists and engineers, managers and licensing officers, attorneys, professional patent searchers, data sellers, patent examiners, journalists, judges. In many cases the core inventors have little input into the process. In our field work, my colleagues and I have interviewed lawyers who tell us that hours may be spent debating the placement of a single word in a claim, and that therein lies the value. Professional patent searchers swap ideas in active professional communities, and attend conventions where they debate the semantics and semiotics of search strategies, all the while worrying about earning a living from hourly fees. Patent examiners take field trips to Silicon Valley to become up-to-date on latest prior art, and engage in lengthy negotiations with applicants after patent applications have been filed. Career paths of patent professionals are, as in the military–industrial complex, highly incestuous, with frequent mobility across the three major fields of specializations: patent examination, prior art search, and the law.

The complex sociology by which patents are made has implications for empirical research using patents. As the samples sizes get larger, our point estimates are made with more precision, but our interpretation remains dependent on an understanding of what the data mean. We should learn from ethnographers who collect data on the conversations, strategies, and interactions that are missing in the patent data. We should take care not to de-construct the different components of these documents and re-combine them in models that claim to show causal relationships. The 90% unexplained variance between our strongest bibliometric measures of value and the monetary value as reported by inventors reminds us that the patent data are not neutral and objective encodings of inventions but may represent interesting and valuable inventions in their own right. The results of this valuable article stand as a challenging invitation to explore these issues further.

Notes

1 See, for example, Levin *et al.* (1987), Trajtenberg (1990), Harhoff *et al.* (1999), Jaffe *et al.* (2000), Cohen *et al.* (2000), Lanjouw and Schankerman (2004), Hall *et al.* (2005), Thompson and Fox Kean (2005), Alcacer and Gittelman (2006), and Thompson (2006).

2 For example, Arora *et al.* (2001).

3 It is interesting that Hall *et al.* (2005) report much higher R^2 in regressions that link patent citations to value of the firm (Tobin's *Q*), and estimate that every additional patent citation adds 3% to firm market value, and that firms with the most highly cited patents are granted a 54% market premium. It is difficult to square those results with the results reported here – especially since the link from citations to patent value should be stronger than between citations and firm value – unless we accept that much of the value contained in citations is largely a market signaling effect for the firm. This interpretation is in line with Lanjouw and Schankerman's (2004) finding that patent quality (composed of the four bibliometric value measures used here) does not correlate well with firm-level research productivity but does impact firm market value.

References

Alcacer, J. and M. Gittelman, 2004, "How do I know what you know? Patent examiners and the generation of patent citations". Available at SSRN: http://ssrn.com/abstract = 548003.

Alcacer, J. and M. Gittelman, 2006, "Patent citations as a measure of knowledge flows: The influence of examiner citations". *Review of Economics and Statistics*, **88**: 774–779.

Alcacer, J., M. Gittelman and B. Sampat, 2008, "Assignee and examiner citations in U.S. patents: An overview". Unpublished manuscript.

Arora, Ashish, Marco Ceccagnoli and Wesley Cohen, 2008, "R&D and the patent premium". *International Journal of Industrial Organization*, doi: 10.1016/j.ijindorg.2007.11.004.

Arora, A., A. Fosfuri and A. Gambardella, 2001, *Markets for technology*. Cambridge, MA: MIT Press.

Cockburn, I., S. Kortum and S. Stern, 2004, "Are all patent examiners equal? Examiners, patent characteristics, and litigation outcomes". in W. M. Cohen and S. Merrill (eds.) *Patents in the knowledge-based economy*. Washington DC: National Academy Press.

Cohen, Wesley M., Richard R. Nelson and John P. Walsh, 2000, "Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)". NBER Working paper 7552 (February).

Hall, Bronwyn, Adam Jaffe and Manuel Trajtenberg, 2005, "Market value and patent citations". *RAND Journal of Economics*, **36**(1): 16–38.

Harhoff, Dietmar, Francis Narin, Frederic Scherer and Katrin Vopel, 1999, "Citation frequency and the value of patented innovation". *Review of Economics and Statistics*, **81**: 511–515.

Hegde, D. and B. N. Sampat, 2005, "Examiner citations, applicant citations, and the private value of patents". Paper presented at the 2005 REER meeting, Georgia Tech.

Jaffe, Adam B., Manuel Trajtenberg and Michael S. Forgarty, 2000, "Knowledge spillovers and patent citations: Evidence from a survey of inventors". *The American Economic Review*, **90**(2): 215–218.

Lanjouw, Jenny and Mark Schankerman, 2004, "Patent quality and research productivity: Measuring innovation with multiple indicators". *Economic Journal*, **114**: 441–465.

Lemley, Mark and C. Shapiro, 2005, "Probabilistic patents". *Journal of Economic Perspectives*, **19**(2): 75–98.

Levin, R. C., A. K. Klevorick, R. R. Nelson. and S. G. Winter, (1987), "Appropriating the returns from industrial research and development". *Brookings Papers on Economic Activity*, **1987**(3): 783–820.

Schneider, C., 2007, "The Determinants of Patent Applications Outcomes – Does Experience Matter?". MPRA Paper No. 3359, http://mpra.ub.uni-muenchen.de/3359/.

Thompson, P., 2006, "Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations". *The Review of Economics and Statistics*, **88**(2): 383–388.

Thompson, P. and M. Fox Kean, 2005, "Patent citations and the geography of knowledge spillovers: A reassessment". *American Economic Review*, **95**(1): 450–460.

Trajtenberg, M., 1990, "A penny for your quotes: Patent citations and the value of innovations". *The RAND Journal of Economics*, **21**: 172–187.