European Management Review (2008) 5, 69–84

palgrave-journals.com/emr

The value of European patents

Alfonso Gambardella[1], Dietmar Harhoff[2], Bart Verspagen[3]

[1]Department of Management, Bocconi University, Milan, Italy;
[2]Inno-Tec, Ludwig-Maximilians-Universität (LMU), München, and CEPR, London;
[3]Department of Economics & UNU-MERIT, Maastricht University, Maastricht, The Netherlands

Correspondence: Alfonso Gambardella, Bocconi University, Milan, Italy.
E-mail: alfonso.gambardella@unibocconi.it

Abstract

This paper employs data from an extensive European survey to produce one of the first systematic assessments of the private economic value of patents. The estimated mean of our patent value distribution is higher than 3 million euros, the median is about one-tenth of it, and the mode is around a few thousand euros. This is in line with previous findings about the skewed distribution of patent values. Our measure is significantly correlated with the number of patent citations, references, claims, and countries in which the patent is applied. Citations explain value as much as the other three indicators combined, and the right tail of citations is correlated with the right tail of our value measure. Yet, the four indicators only explain 2.7% of the variance of patent value. Thus, while the use of these indicators as proxies for value, particularly citations, may be justified, predictions based on these indicators carry significant noise. After using country, technology, and patent class fixed effects, we only explain 11.3% of the variation in patent value. The 'measure of our ignorance' about the determinants of patent value is then still sizable, which calls for additional research to fill the gap.

European Management Review (2008) **5,** 69–84. doi:10.1057/emr.2008.10

Keywords: patents; intellectual property rights; value of patents; patent citations; patent value indicators; patent references; patent claims

Introduction

The search for reliable estimates of the economic value of patents has received considerable attention among economists, business scholars, and policy makers.[1] This is parallel to an increase in the relevance of intangibles (including inventions and know-how) for firm value over the last two decades. Moreover, as the number of patent applications has surged in Europe, Japan and the US, economists and management scholars have become more and more dissatisfied with using simple application or grant numbers as an indication of R&D output.

Against this background, this paper estimates the economic value of patents by employing a unique and comprehensive data set drawn from a large scale survey of European inventors. The PatVal-EU survey collected data on more than 9000 patents (out of more than 27,000 questionnaire mailings), including their value and a broad set of characteristics describing the context of the invention. These are patents with priority date 1993–1997 granted by the European Patent Office (EPO), and such that the address of the first inventor listed in the patent is in Denmark, France, Germany, Hungary, Italy, the Netherlands, Spain or the UK. The survey data are obtained from questionnaire responses produced by the first inventor or, if the first inventor was not available, by any other inventor on the patent whose address is in one of our eight countries.

Our contribution is twofold. First, we measure and estimate the economic value of patents. Second, we assess the relationships between our survey-based measure and some indirect indicators commonly used in the literature, viz. the number of forward citations, backward references, claims, and countries in which the patent is applied for. The first contribution offers a benchmark for assessing the value of patents as a firm asset. At present, there is little evidence about the economic value of patents. In light of the increase in the number of patents since the 1980s, this means that we may be ignoring, or poorly estimating, a significant portion of the assets of an economy. In addition, estimates of the economic value of patents help assess the firm's portfolio of intellectual property or proprietary knowledge, which is increasingly important. Also, there are increasingly firms, typically smaller companies and start-ups, that rely only on intangible capital – usually ideas or patents. More precise valuation of these companies (for VC investments, lending decisions or financial valuation at the time of an IPO or trade sale) will be highly beneficial as it may lower the cost of capital; but that will rely to a considerable degree on improved techniques for assessing the value of patents.

Our second contribution speaks to the many studies that use patents in the management and economics literature. Apart from their use to measure knowledge flows (e.g., Almeida and Kogut, 1999; Alcacer and Gittelman, 2006), patents are used as a proxy for innovation performance. However, this is bound by the limitations of patent counts, which do not account for quality or heterogeneity across patents. The literature has then resorted to the use of forward citations, or other indicators, as proxies for innovation quality or economic value (e.g., Trajtenberg, 1990; Gittelman and Kogut, 2003; Lanjouw and Schankerman, 2004; Singh, 2008). Yet, the limited availability of direct measures of value implies that the relationships between these indicators and the value of patents are still largely untested. Among the few studies that test these relationships, Hall *et al.* (2005) find that the stock of patent citations of the firms are significantly correlated with their market value. Compared to Hall *et al.* we test these relationships at the level of the individual patent, we study the impact of citations along with our three other indicators, and we assess the extent of their individual or joint contribution in explaining value.

Our measure of the value of patents follows Harhoff *et al.* (2003a), and it is given by the answer to the following question to the inventor named on the patent document: 'Suppose that on the day on which this patent was granted, the applicant had all the information about the value of the patent that is available today. In case a potential competitor of the applicant was interested in buying the patent, what would be the minimum price (in euro) the applicant should demand?' We offer a menu of 10 interval responses: less than €30K; 30–100K; 100–300K; 300K–1M; 1–3M; 3–10M; 10–30M; 30–100M; 100–300M; more than 300M. Because the question asks about a hypothetical situation in which the patent is sold, the value measure obtained from the answer logically includes a strategic component of patent value not captured by the renewal value. The reason is that the buyer of the patent will obtain the exclusive right to the patent and may be able to block related patents of the previous owner, thus preventing the original patent holder from practicing these inventions or demand license fees for them. Another way to think about our measure is that it is equivalent to the market value of a firm whose only asset is the patent.

This paper first studies the distribution of the measure of patent value obtained from the PatVal-EU survey to assess whether its shape conforms to what we know about patent value distributions, and whether the magnitudes involved are reasonable. It then addresses the problem that the inventors may be less informed than managers about the value of their patents. For a sample of French patents, the value question was submitted to both the inventor and a manager responsible for the development of the patent. We find that the inventors slightly overestimate the value of their patents. However, the bias is negligible.

In order to estimate the economic value of patents we regress the log of the mid-points of our value classes on the four indicators above, and on country, application-year, and technology or patent class fixed effects. We use weights to take into account selection biases in the PatVal-EU observations compared to the universe of patents. We find that the estimated mean-value of our patent distribution is about 3 million euros, while the median is slightly lower than 400 thousand euros. Discrepancies of this size between median and mean are not untypical for highly skewed distributions. The mean is affected by the long tails of the distribution, and it is sensitive to the extreme values of the support. In some theoretical distributions with these characteristics the mean or some higher moments do not even exist. In our distribution we find that even the mode is smaller than the median.

We find that citations explain patent value as much as the other three indicators. Moreover, citations mirror the right tail of the patent distribution: other things being equal, patents in the top 5 or 1% of the distribution of citations are significantly more likely to be in our very top value classes. Our four indicators combined explain only 2.7% of the variance in our value measure. This suggests that our sampling weight methodology and our empirical analysis more generally adequately represents the universe of EPO patents.

The remainder of the paper proceeds as follows. The second section discusses the existing literature on measuring patent value as a benchmark for assessing the patent value from our survey. The third section discusses the variables employed in our analysis and their relationship to the theoretical literature. The fourth section presents a validation of our survey measure of patent value. The fifth section presents our empirical framework and results as well as estimates for the mean and other important statistics of the patent value distribution. The sixth section studies the relationships between patent value and the indirect indicators, particularly citations. The final section concludes.

Background studies on patent value

The initial approaches to measuring the value of patents have relied on data on patent renewal. The obligation to pay renewal fees to keep patents 'alive' implies that it is expensive to patent holders to renew patent protection for an additional year. The pioneering papers in this field are Pakes (1986), and Schankerman and Pakes (1986). A more recent contribution that covers a large set of explanatory variables is Bessen (2006). This approach relies on the fact that a substantial amount of patents is not renewed until the end of the statutory life time (in Schankerman and Pakes (1986) only around 10% of the patents issued in Germany, France and the UK 'lives' for the full statutory term). Based on such an approach, Schankerman and Pakes (1986) report estimates of the means and medians of the patent value distribution for patents issued in 1970 in Germany, France, and the UK. The median values in these countries are, respectively, \$17,329, \$847 and \$1861 (all in 1980 prices). The distribution means are substantially higher than the median in France and the UK (\$6656 and \$6963, respectively), and slightly higher in Germany (\$19,124). This is indicative of the skewed nature of the data.

However, there are reasons to believe that the renewal fee approach does not accurately capture the value of patents that are in the extreme right tail of the distribution. For patents renewed over their statutory lifetime, the renewal fees provide a lower rather than an upper bound for patent value. The renewal approach also assumes that the annual returns from having the patent in force decrease monotonically over the patent's life. An 'early bloomer' effect cannot be detected. Thus, even the patents that are discontinued before they reach the maximum term may have generated significant returns in the very first years of their life before they are discontinued. In short, the distribution of patent values may be more skewed than predicted by the renewal approach, and patent values may be underestimated for that reason.

In addition, the renewal data-based estimates provide information only on the part of patent value that Arora *et al.* (2008) call 'patent premium.' These authors make a distinction between the value of the invention to the firm without a patent being issued, and the extra value that the patent generates to the firm (premium). The relevance of this distinction arises from the finding by, among others, Levin *et al.* (1987), that a substantial number of inventions are not patented, but are protected by other means (e.g., secrecy and lead time). Harhoff *et al.* (2003b) show that the renewal decision will take into account that if a patent is not renewed, the invention may still be practiced by the former owner. If the patent is sold, however, as we assume in our survey question about value, the new owner may block revenues from complementary inventions as well. Harhoff *et al.* (2003b) distinguish between the 'renewal value' and the 'asset value' of the patent right. Both are two conceptually different types of the patent premium discussed by Arora *et al.* (2008). The renewal value will again tend to underestimate the asset value, since it ignores the strategic role of the exclusion right in the context of cumulative or complementary inventions.

Serrano (2005) uses data on the commercial transfer of patents rights, as registered at the USPTO, as an estimator of patent value. He applies a model that resembles the Pakes (1986) model, in that it follows the value of the patent over its lifetime. In addition to the decision to renew the patent, however, the Serrano model also includes an option to sell the patent. Using estimated model parameters, Serrano estimates the median value in his sample of patents to be equal to $27,895 and the mean equal to $86,782 (both 2003 prices). These values include patent premium and invention value, that is, they represent total patent value. Note, however, that the sample for which the model is estimated contains only patents applied for by 'small innovators' (firms that have no more than five patents granted per year).

Patent value may also be inferred from econometric models that link firms' market value to patents (Griliches, 1981; Pakes, 1985). Hall *et al.* (2005) use data on Compustat firms during 1979–1988 and show that the firms' market value is correlated with the ratio between R&D and firm assets, the ratio between patents and R&D, and the ratio between citations and patents. They find that a 1% increase in R&D on assets, or in patents on R&D, or an extra citation produce increases in the market value of the firm in the order of million dollars. Bessen (2007) argues that the regression coefficient on patents in an estimated market value equation also captures an effect related to the quality of R&D (firms that perform high-quality R&D are identified through a higher patents-to-R&D ratio), and hence this coefficient is an upper bound rather than a precise estimator for the value of patents. He estimates the upper bound for a sample of US firms for the period between 1969 and 2001, and arrives at an estimate of 376,000 US$ (1992 prices) as the average patent value.

A more indirect approach to approximating patent value has been to use proxy variables, such as citations, and more recently, in the European setting, the filing of a legal opposition to the patents (Harhoff *et al.*, 2003b). Forward citations account for the visibility and importance of the patent. As Trajtenberg (1990) has shown, citation measures are correlated with a patent social value. Given the costs of legal battles, only privately valuable patents are worth opposing, as shown theoretically by Harhoff and Reitzig (2004). Lanjouw and Schankerman (2004) develop a combined index that uses a set of indirect measures to infer patent value from the correlation structure of observable patent characteristics, but does not build on observed patent value data. As noted above, Hall *et al.* (2005) also find that citations are quite valuable.

Our approach to measuring patent value treats patents as an asset, and in asking for the price at which the asset would be sold at the moment of grant it tries to retrieve the inventor's assessment of the discounted flow of profits accruing to the patent applicant because of the patent. In so doing, our value indicator includes both the invention value and the patent premium. It will therefore yield higher estimates of patent value than an approach based on renewal data, since the latter, as argued above, only estimates the patent premium part of the total value. Moreover, for broad patents that cover key features of products or processes – a case found in interviews to be common for the most valuable patents – the sale of rights puts at risk the whole stream of quasi-rents realizable from a product or process. Or at minimum, the sale of rights can impose upon the seller the profit sacrifice from foregoing key patented features or the cost of inventing around them. Hence we may expect that our patent values estimates can be quite substantial at the higher end of the distribution (hence the 300 million euro right cut-off). We discuss further details of our value estimator in the next section.

The patent value measure from PatVal-EU

The PatVal-EU survey, from which the patent value data of this paper are obtained, is an extensive survey conducted in 2003–2004 on EPO patents with priority date 1993–1997 and whose first inventor was in one of the eight countries listed in the introduction. The basic PatVal-EU paper (Giuri *et al.*, 2007) provides the relevant information about how the survey was conducted, its underlying issues and problems, and it shows the basic descriptive statistics from the questionnaire.[2]

Table 1 defines all the variables that we employ in this analysis. Table 2 presents descriptive statistics. In this paper we employ the 8217 PatVal-EU observations with valid responses to our question about the value of the patent. Figure 1 reports the distribution of these answers. The distribution is skewed to the left, and it conforms to other assessments of the value of patents in the literature

Table 1 Description of variables employed in the analysis

Variable	Description
VALUE	Index equal to 1–10 for the following PatVal-EU classes of patent values: ⩽€30K; 30–100K; 100–300K; 300K–1M; 1–3M; 3–10M; 10–30M; 30–100M; 100–300M; ⩾300M
VALUEM	Mid point of VALUE (15K; 65K; 200K; 650K; 2M; 6.5M; 20M; 65M; 200M; 650M[a])
CITES	# of forward citations to the patent within 5 years after the publication of the patent (usually 18 months after the priority date), including citations to equivalent patents
REFS	# of backward references in the patent
CLAIMS	# of claims of the patent at the moment of grant
STATES	# of designated European countries in which the patent is applied for
CITES0–5	Six dummies for CITES = 0; 1; 2; 3–5; 6–8; or 9, corresponding to the following percentiles of the CITES distribution of all the EPO patents with priority date 1993–1997 granted by 2003 and with first inventor in our eight countries (49,941 patents): 1–45; 46–70; 71–83; 84–96; 96–98; ⩾99.
VALUE⩾5, VALUE⩾6, VALUE⩾7	Three dummies equal to 1 if VALUE ⩾5, 6 or 7, corresponding to 17.5, 7.8, or 3.9% of the 8217 PatVal-EU patents for which data on VALUE are available.
Country dummies	Eight dummies for address of the first inventor in Denmark, France, Germany, Hungary, Italy, Netherlands, Spain, UK
Application year dummies	Six dummies for application years 1993–1998[b].
Technology dummies	Thirty technological area dummies obtained by converting the IPC classes of the patent using the ISI-INPI-OST concordance list[c].
IPC 3-digit dummies	117 dummies for the main IPC 3-digit class of the patent

[a]For the last interval of VALUE, VALUEM was set equal to the mid-point of 300–1000M.
[b]PatVal-EU sampled patents with priority date 1993–1997, but applications may exhibit a later date because applicants may patent first in their national countries. Hence, for some PatVal-EU patents application year is 1998.
[c]See Giuri *et al.* (2007) and Schmoch and Kirsch (1993) for details and references on the ISI-INPI OST concordance list. The 30 technology areas are: Agricultural and food processing, machinery and apparatus; Agriculture, food chemistry; Analysis, measurement, control technology; Audio-visual technology; Biotechnology; Chemical and petrol industry, basic materials chemistry; Chemical engineering; Civil engineering, building, mining; Consumer goods and equipment; Electrical devices, electrical engineering, electrical energy; Engines, pumps, turbines; Environmental technology; Handling, printing; Information technology; Machine tools; Macromolecular chemistry, polymers; Materials processing, textiles, paper; Materials, metallurgy; Mechanical Elements; Medical technology; Nuclear engineering; Optics; Organic fine chemistry; Pharmaceuticals, cosmetics; Semiconductors; Space technology weapons; Surface technology, coating; Telecommunications; Thermal processes and apparatus; Transport.

(Harhoff *et al.*, 1999; Scherer and Harhoff, 2000; Scherer *et al.*, 2000).

As Table 2 shows, the sample average of VALUEM, the mid-point of the value intervals, is higher than 10 million euros. The median is 650 thousand euros. Compared to the value estimates that we quoted in the second section, these estimates are an order of magnitude larger than what is usually obtained. This is especially true of the averages (as opposed to the median value). In relation to the value estimates based on renewal data, this higher value is related to the fact that the renewal-based estimates only include patent premium. But our results are still an order of magnitude higher than the highest estimate of patent value that we saw in the second section, that is, Bessen's (2007) estimate of 376,000 1992 US$, based on a methodology that includes patent premium and invention value.

The high average is strongly influenced by the few patents in the extreme right classes. For example, if we exclude the 75 observations in the highest value class, the sample average for VALUEM drops to 5.0 million euros. Further dropping the 68 observation in the one-but-highest class brings down the average to 3.4 million euros (the median values are not affected by these reductions in the sample). These calculations are in line with the conclusions from the literature that looks at patent value distributions using extreme value theory (e.g., Silverberg and Verspagen, 2007).

This particular characteristic of the distribution may indeed cause previous estimates of mean patent values to be underestimated. Serrano's estimates depend on traded patents, and if the most valuable patents are not traded, or they are in the hands of large corporations with more than five patents per year, this would seriously bias downwards the average value that he finds. In regressions linking firms' market values to patents (e.g., Bessen, 2007), the extreme value observations will appear as outliers with only a marginal effect on the estimated regression parameters (and hence average patent value). Although we feel that when viewed in this way, our results are not necessarily inconsistent with previous estimates of patent value in the literature, we need to put our estimates under scrutiny also independently of the previous estimates of patent value.

An important question in this respect is whether the scale that we used in the questionnaire is reasonable. We have several reasons to believe that it is. First, from Figure 1 less than 1% of patents for which we obtained a response had a declared value higher than 300 million euros (our highest class). We checked these patents and while we cannot completely rule out odd answers, quite a few of them

Table 2 Descriptive statistics

	Mean	St. Dev.	Min	P25	Median	P75	Max	N .obs.
VALUE	3.864	1.831	1	3	4	5	10	8217
VALUEM	10887.14	64647.08	15	200	650	2000	650,000	8217
CITES	1.449	2.228	0	0	1	2	40	8217
REFS	4.38	2.241	0	3	4	6	18	8217
CLAIMS	10.957	7.27	1	6	10	14	131	8217
STATES	9.009	4.92	1	5	7	13	19	8217
DE	0.373	0.484	0	0	0	1	1	8217
DK	0.052	0.223	0	0	0	0	1	8217
ES	0.016	0.126	0	0	0	0	1	8217
FR	0.132	0.339	0	0	0	0	1	8217
HU	0.004	0.065	0	0	0	0	1	8217
IT	0.128	0.334	0	0	0	0	1	8217
NL	0.127	0.333	0	0	0	0	1	8217
UK	0.166	0.372	0	0	0	0	1	8217
Applic. year 93	0.025	0.157	0	0	0	0	1	8217
Applic. year 94	0.28	0.449	0	0	0	1	1	8217
Applic. year 95	0.254	0.436	0	0	0	1	1	8217
Applic. year 96	0.228	0.419	0	0	0	0	1	8217
Applic. year 97	0.159	0.366	0	0	0	0	1	8217
Applic. year 98	0.054	0.225	0	0	0	0	1	8217

Descriptive statistics computed for the 8217 PatVal-EU patents for which VALUE is available. Descriptive statistics for CITES1–5 and for VALUE$\geqslant$5–7 are straightforward. The 8217 PatVal-EU patents employed in our analysis are spread fairly well across the industry and IPC three-digit dummies.

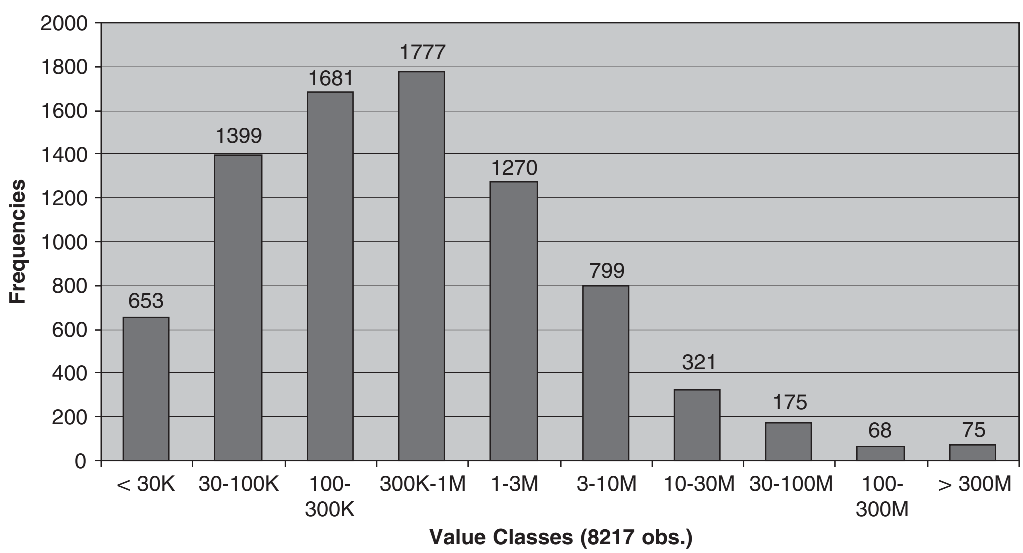

Figure 1 Distribution of VALUE. The figure shows that the PatVal-EU patent VALUE distribution is skewed. Since the difference in the logs of the boundaries of the intervals is roughly constant, the distribution in the figure is an approximation of a log-normal. Even the log-normal distribution looks skewed.

regarded pharmaceutical products, which are generally quite valuable, or other major product innovations. Second, is a patent value higher than 300 million euros really abnormal? Suppose that a patent provides a monopoly power on a product for about 20 years. This is roughly the length of patent life plus some adjustment years for competition to pick this up. Assuming 5% discount rate and a constant flow of profits from this asset, simple calculations show that an asset worth 500 million euros commands a constant annual flow of profits of slightly less than 40 million euros. Some pharmaceutical products have annual sales in the order of hundreds of millions of euros, and a 40 million euro rent seems a conservative order of magnitude of the profits of a very small share of highly valuable patents.

A reason why there may be a potential upward bias of our measure is that respondents may inflate the reported patent values in order to boost their perceived performance, or they may be reluctant to state that their patents have zero value, or they are more reluctant to respond if the patent that we asked about has a low value. To account for the first two of these potential tendencies, we set VALUEM equal to the midpoint of the previous interval (i.e., 0 for the first interval, 15,000 for the second, 65,000 for the third, etc.). We consider this to be a way to bias our estimator downward in a quite conservative way. If we also exclude the 143 patents in the two highest value classes, the mean of the distribution of VALUEM drops to about 1 million euros. This is still higher than the Bessen estimate, and hence we conclude that our data suggest a higher average patent value than the studies reviewed in the second section.[3]

Our measure may also include strategic elements in the utility of the patent holder, like the desire of preventing others from using one's own invention or similar non-economic attitudes. A patent holder who is reluctant to give out the patent right may state a high value. Also, our measure is not the actual realization of a market transaction involving the patent, as there may be no buyers of the patent at that price. Yet, if the owner is not willing to part herself from the patent at a lower price, it means that she may be able to make at least that much, or more generally that this is her reservation price, which for the reasons just noted may encompass both monetary and non-monetary aspects.

We can provide some control on the questionnaire responses because German inventors can be expected to have a more precise idea of the economic value of their patents. The German Employees Inventor Compensation Act establishes that German employers can claim the patent rights from an inventor by providing him with a fair compensation, which provides the German inventors with a good reference point for their PatVal-EU answers (Harhoff and Hoisl, 2007). A specific PatVal-EU question asked whether the inventors were compensated for their inventions, and 62% of German inventors in our sample were compensated for their patents whilst that share was well below 30% for all other countries. This suggests that German inventors have more precise information of the economic value of their patents. The sample average of VALUEM for the German patents is 5.6 million euros, which is slightly more than half of the 10.9 average of the full sample reported in Table 2. Moreover, the German median is 200.000 euros, i.e., in the class just before the median class for the full sample VALUEM. Although this comparison may be influenced by variations in the relative frequency of the most valuable patents in the German and total sample, these numbers confirm that German inventors are more conservative in their assessments, although not an order of magnitude more conservative.

Finally, we can compare our value estimator to the one obtained by other studies that used similar methodologies. Scherer and Harhoff (2000) sample 772 patents filed in Germany in 1977, and produced an average value of about 5 million Deutsche Marks, or about 2.5 million euros (1977 currency). Their patents are 16–20 years older than the ones in our data assuming a 5% inflation rate this amounts to about 6 million euros 18 years later, which is very close to the German average in PatVal-EU of 5.6 million euros. Also, the median value of the 772 German patents in Scherer and Harhoff is the value class 100–400K Deutsche Marks, or about 50–200K euros. Again, given that their patents are on average 18 years older than the PatVal-EU patents, and assuming 5% inflation, the midpoint of 125K euros becomes about 300K, which is quite close to the German PatVal-EU median of 200K euros.

Comparing inventor and manager responses

A potential limitation of our measure of patent value is that it is reported by inventors. Especially in large firms, or even in academic settings, managers may provide more accurate estimates of the value of a patent. The trade-off is that if one wants to conduct a survey at the scale of PatVal-EU, it is costly to seek for each patent the most suited individual to answer such a question. Moreover, since we are dealing with patents that are some years old, these individuals might have left the company. Thus, even if inventors may offer less precise answers, it is not at all clear that we would not have introduced more serious biases by seeking other respondents to the value question. Inventors are likely to be the easiest to reach and reasonably knowledgeable individuals who know about the patent and can provide a 'good' guess systematically and on a large scale.

At any rate, for a sample of 354 French patents the question about the value was asked independently to the inventor and a manager. Figure 2 shows the distributions of the two respondent classes. Figure 3 shows the distribution of the difference between the 1 and 10 number of the class picked by the inventor and the manager. The two distributions in Figure 2 overlap to a great extent. Figure 3 shows that in slightly more than two-third of the patents the inventors and managers missed each other by at most one contiguous class (difference between -1 and 1), and for almost 90% of the patents they missed each other by at most two contiguous classes (-2; 2).

Tables 3 and 4 compare the two distributions more formally. Table 3 shows that the inventors report a higher mean response of the 1–10 VALUE index than managers. Table 4 reports statistical tests. It shows that a two-tailed t-test of differences in the mean responses does not reject the null-hypothesis of equal means at significance levels smaller than 8%. In fact, while the inventors may boost the results of their work, it is harder to think that the managers may over-estimate the value of patents. It may then be reasonable to employ a one-tailed t-test of the null hypothesis against the alternative that the mean response of the inventors is higher than that of the managers. Table 4 shows that in this case the null hypothesis of equality of the means is rejected at significance level $<4\%$. Table 4 also reports other tests. In all of them we never reject the null hypothesis of equal responses between inventors and managers. In particular, we cannot reject the hypothesis of equality of the standard deviations of the two distributions, and the Kolmogorov–Smirnov and Wilcoxon rank-sum (Mann–Whitney) test do not reject the hypothesis that the two distributions are equal. In sum, our results show that the inventors slightly overestimate the economic value of their patents but that this overestimation is not particularly severe.

Compared to smaller firms or other organizations (universities, research labs) the inventors in large companies may be less informed about the value of their patents because of the greater organizational distance and the more intensive specialization of tasks. As a result, the gap in response ought to be wider. Table 5 corroborates this hypothesis. Inventors in large firms exhibit a higher

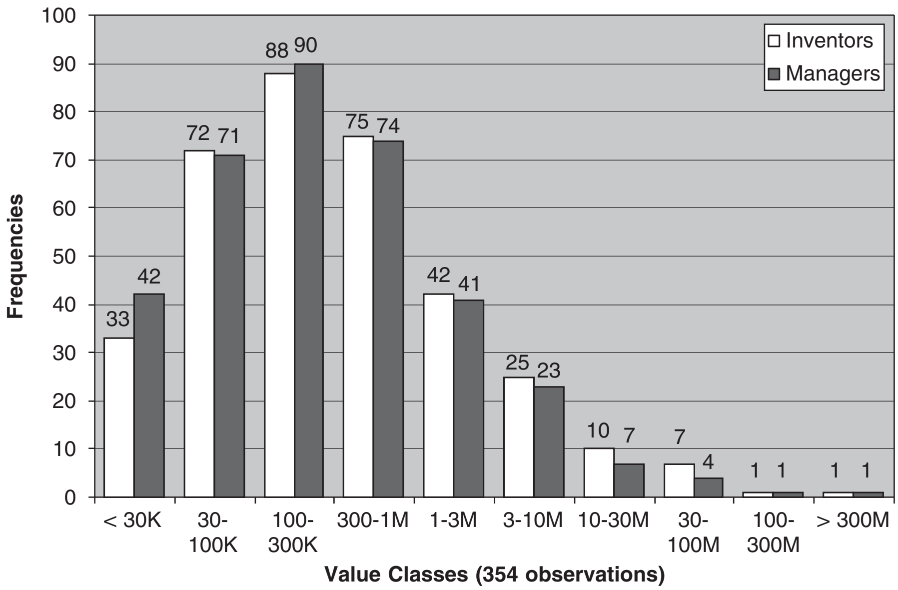

Figure 2 Distribution of VALUE, responses by 354 French inventors and managers. VALUE responses by 354 French inventors and managers who were responsible for the patent and provided independent responses about its value of the same patent. The figure shows that the two distributions are similar.

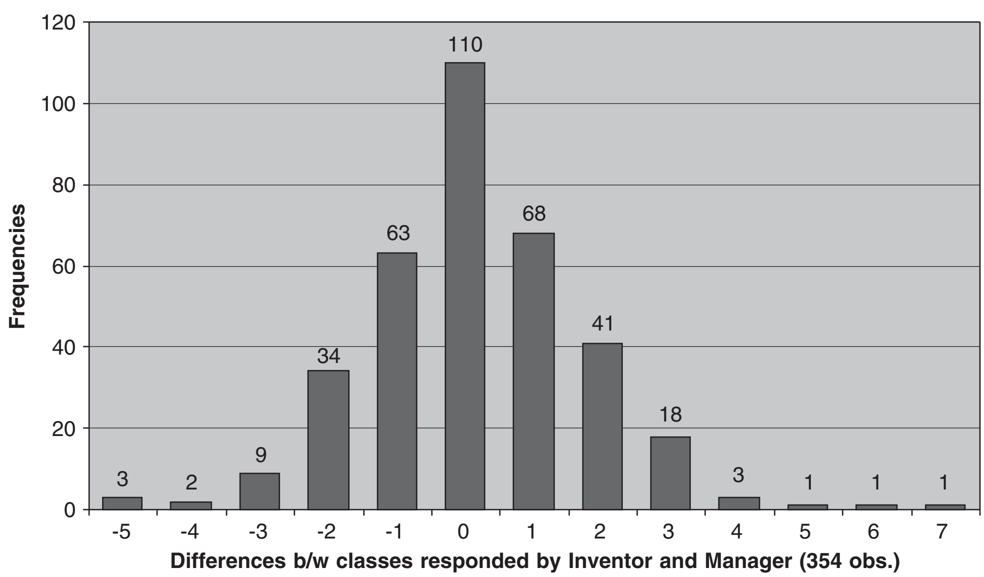

Figure 3 Differences in VALUE, responses by 354 French inventors and managers. VALUE responses by 354 French inventors and managers who were responsible for the patent and provided independent responses about its value. The figure shows that almost 90% of the responses fall within two VALUE classes.

average difference in their assessment of patent value with respect to their managers in other organizations. Table 6 further investigates the difference between large firms and the rest of the sample. It first shows that the equality of mean responses between inventors and managers is rejected for the large firms (two-tailed at significance $<10\%$, one tailed at significance $<5\%$), while it cannot be rejected for the other organizations. In addition, one cannot reject the hypothesis that the average difference in the inventor-manager responses in large firms are equal to other organizations, and one cannot either reject the hypothesis that the standard deviations of the two distributions of the differences are equal. Finally, one cannot reject the hypothesis of the equality of the distributions of the differences according to the Kolmogorv–Smirnov and the Wilcoxon rank-sum (Mann–Whitney) test. While the lack of significance of these tests may stem in part from their small number of observations, it is also a consequence of the fact that the differences are probably not highly pronounced.

To summarize, the slight overestimate of the inventor assessment of the value of their patents compared to managers seems to be produced by inventors in large firms. This also helps to qualify our earlier remark that inventors in smaller firms or other organizations may be more likely to be biased in their assessment of patent values. Our results suggest that this is not the case, and that their evaluations are even closer to those of their managers. Yet,

Table 3 Comparing the responses to the value question of French inventors and managers, VALUE classes 1–10 (354 obs.)

VALUE reported by	*Mean*	*Std. error*	*Std. Dev.*	*Min*	*P25*	*Median*	*P75*	*Max*
Inventors	3.520	0.089	1.680	1	2	3	4	10
Managers	3.370	0.086	1.625	1	2	3	4	10
Difference	0.150	0.086	1.608	−5	−1	0	1	7

Means and other statistics of VALUE for the 354 French PatVal-EU patents with an answer for VALUE from both a manager and an inventor. *Std. Dev.* is the standard deviation of the VALUE responses. *Std. error* = *Std. Dev.*/(354)$^{1/2}$, that is, it is the estimated standard deviation of the mean. The table shows that the mean of VALUE for the inventor responses is only slightly higher than that of the managers. This suggests that the inventors only slightly overestimate patent value compared to managers.

Table 4 Means, standard deviations, distributions

Test	*P-value*
t-test for difference between the means of VALUE in inventor vs manager responses (H_0: Mean diff. = 0)[a]	
• Two-tailed test	0.084[b]
• One-tailed test (mean inventors > mean managers)	0.040[c]
Two tail F-test for difference between Std.Dev. (H_0: Diff. in Std. Dev. = 0)[d]	0.534
Two sample Kolmogorov–Smirnov test for equality of distributions[e]	0.754
Two-sample Wilcoxon rank-sum (Mann–Whitney) test for equality of distributions[e]	0.286

Tests of differences in the responses of French inventors and managers, VALUE classes 1–10 (354 obs.)

[a]Means of VALUE for inventor and manager responses from Table 3.

[b]Null hypothesis rejected at $P<10\%$.

[c]Null hypothesis rejected at $P<5\%$.

[d]Std. Dev.s of VALUE for inventor and manager responses from Table 3.

[e]Distributions of VALUE for inventor and manager responses.

The table shows that the difference in the mean responses of VALUE between inventors and managers in the previous table are statistically significant. It also shows that the standard deviations and the distributions of the VALUE responses of inventors and managers are not statistically different.

Table 5 Differences across organizations in the responses of French inventors and managers, VALUE classes 1–10 (350 obs.[a])

Difference in VALUE by (# obs.)[b]	*Mean*	*Std. error*	*Std. Dev.*	*Min*	*P25*	*Median*	*P75*	*Max*
Large firms (207)[c]	0.188	0.113	1.630	−5	−1	0	1	6
All others (143)[d]	0.077	0.136	1.628	−5	−1	0	1	7

[a]Organization type could not be attributed in four observations.

[b]Difference in the VALUE responses of inventors and managers by large firms or all other organizations.

[c]Large firms = firms with > 250 employees.

[d]All others = smaller firms, universities, other organizations.

The table shows that the mean difference in VALUE between inventors and managers is higher in large firms compared to other organizations.

in these other organizations the managers are themselves more directly and closely involved with the invention and they themselves may be biased in their evaluations. For example, in a small start-up the inventor and manager probably work closely together. Similarly in a university researchers and managers of the technology transfer office discuss a great deal about the potential value of the patent. Again, we cannot rule out this hypothesis. However, the average values of the 1–10 numbers for the value classes chosen by inventor and managers in the case of large firms are, respectively, 3.57 and 3.39 as compared to 3.41 and 3.34 for all other organizations. Thus, the absolute value of the manager evaluation in other organizations is even lower than in large firms, which suggests that on average managers in small firms or universities do not overestimate the value of patents *vis-à-vis* managers in large firms. Of course, these averages do not control for several other factors that may affect the expected value of patents in different organizations. However, our inspection of the data shows that they do not entail a substantial overvaluation of the value classes in smaller firms or non-profit research institutions as compared to the value measures obtained in firms.

Regression results and estimated distribution of patent values

Regression results

Table 7 reports a significant correlation between our VALUE indicator and the four indirect indicators that we employ in our analysis – cites, references, claims, and countries of application. We show three regressions. In all of them the covariates are our four indicators, along with country, application year, and technology area dummies. The first equation is an OLS regression with log(VALUEM)

Table 6 Tests for differences in the responses by French inventors and managers by organization types, VALUE classes 1–10 (350 obs.[a])

Test	*P-value*
t-test for zero difference between inventor and manager VALUE responses, large firms vs others (H_0: mean diff. = 0)	
Large firms (207 obs.)	
• Two-tailed test	0.098[b]
• One-tailed test (mean inventors > mean managers)	0.049[c]
All others (143 obs.)	
• Two-tailed test	0.573
• One-tailed test (mean inventors > mean managers)	0.286
Two tailed *t*-test for equal difference in inventor-manager mean VALUE responses between large firms and all other organization types (H_0: mean diff. for large firms = mean diff. for all others)	0.530
Two tailed F-test for equal standard deviations of the distributions of the differences in inventor-manager VALUE responses by large firms and others (H_0: st. dev. of diff. for large firms = st. dev. of diff. for all others)	0.989
Two sample Kolmogorov–Smirnov test for equality of the distributions of the differences in inventor-manager VALUE responses for large firms and all others	0.992
Two-sample Wilcoxon rank-sum (Mann–Whitney) test for equality of the distributions of the differences in inventor-manager VALUE responses for large firms and all others	0.475

[a]Organization type could not be attributed in four observations. Large firms and other organizations defined as in Table 5.
[b]Null hypothesis rejected at $P<10\%$.
[c]Null hypothesis rejected at $P<5\%$.
The table shows that mean difference of VALUE between inventors and managers is statistically significant only for large firms. It also shows that: (a) the mean inventor-manager differences in VALUE for large firms and other organizations are not statistically different; (b) the two standard deviations and distributions of the inventor-manager differences in VALUE are not statistically different.

as the dependent variable. If we assume that the error of this equation is normally distributed, the distribution of our value indicator will be log-normal. The second equation is an interval regression. This is an ordered probit with known constants, where the constants are the boundaries of the intervals of our VALUE indicator. Compared to a standard ordered probit we can estimate the standard error of the distribution, which can be retrieved from the estimation given that the constant terms are observed. We measure the boundaries of VALUE in logs, that is log(1)–log(30), log(30)–log(100), and so on. This regression is therefore directly comparable to the OLS regression in the first column because both assume that the dependent variable is the log of the value of the patents. The interval regression, which assumes normality of the errors, also implies that the distribution of value is log-normal.

Finally, in the third column of Table 7 we estimate an OLS regression with the log of the log of VALUEM, that is, log(log(VALUEM)), as the dependent variable. The log-normality assumption has been criticized in the case of the value of patents. Scherer and Harhoff (2000), Harhoff *et al.* (2003a), or Silverberg and Verspagen (2007) all find that the distribution of patent values has a fatter tail than the log-normal distribution, suggesting that there are more highly valued patents at the extreme right tail than predicted by this distribution. With the log of the log we can assume that log(VALUEM) rather than VALUEM is log-normally distributed, which entails that the distribution of log(VALUEM) is skewed rather symmetric. Note that this is consistent with the shape of the distribution of patent values in Figure 1. Because the ratios of the boundaries of the intervals of VALUE are roughly constant, Figure 1, in which the length of the intervals of the classes are equal, depicts a distribution of the log of the variable, and in the figure this distribution looks skewed.

All our regressions in Table 7 are corrected with sampling weights for potential biases in the PatVal-EU sampling.[4] In addition, we employ robust standard errors that we cluster by firm to account for the possibility that patents of the same firms may have correlated errors. The clustering is by ultimate parent firms, which we obtained by searching for the ultimate parent companies of all the firms in the PatVal-EU sample by using *Who Owns Whom* and other company directories.

The three regressions in Table 7 produce very similar results. Since the first and second columns estimate the same specification by OLS or an interval-regression model, the similarity of results implies that they are robust to the use of the VALUE index or its mid-point VALUEM. All four indirect indicators have a statistically significant impact. Forward citations and the designated states in which the patent is applied for have the most important effect.

In all three equations, the German dummy denotes significantly less valuable patents. This is consistent with the conjecture that the German inventors may have some monetary reference to hang their answers about the value of their patents because of the German Employees Inventor Compensation Act. We will then apply the German constant to all the patents to control for potential subjective inflation of patent values in the inventor responses. In all the other countries, but France, the country dummies are higher than Germany. Since we have no special justification for the

Table 7 OLS and interval regressions of patent value

	Dependent Variable log(VALUEM) (OLS)	*Dependent Variable VALUE (Interval regression)*	*Dependent Variable log(log(VALUEM)) (OLS)*
CONST	5.066***	4.834***	1.573***
	(0.000)	(0.000)	(0.000)
LOG(1+CITES)	0.343***	0.353***	0.056***
	(0.000)	(0.000)	(0.000)
LOG(1+REFS)	0.158**	0.163**	0.026**
	(0.010)	(0.010)	(0.013)
LOG(CLAIMS)	0.171***	0.175***	0.030***
	(0.000)	(0.000)	(0.000)
LOG(STATES)	0.372***	0.387***	0.062***
	(0.000)	(0.000)	(0.000)
DE	−0.837***	−0.855***	−0.136***
	(0.000)	(0.000)	(0.000)
DK	−0.148	−0.159	−0.048**
	(0.320)	(0.296)	(0.041)
ES	0.326	0.332	0.048
	(0.128)	(0.128)	(0.141)
FR	−1.05***	−1.088***	−0.184***
	(0.000)	(0.000)	(0.000)
HU	0.141	0.147	0.027
	(0.697)	(0.690)	(0.595)
IT	−0.294***	−0.291***	−0.039**
	(0.003)	(0.003)	(0.012)
NL	−0.303***	−0.312***	−0.055***
	(0.009)	(0.008)	(0.004)
log(σ)		0.720	
		(0.000)	
R^2	0.092	—	0.089
Log-Lik function	—	−7.74E+04	—
N. Observations	8217	8217	8217

P-values based on robust standard errors in parentheses. *$P<10\%$; **$P<5\%$; ***$P<1\%$. All regressions include application year and industry dummies (one omitted), sampling weights, and observations are clustered by patent applicants. The parameter σ is the standard error of the latent variable in the interval regression. The table shows that patent value is correlated with indirect indicators, and this is robust to the type of estimation and to the specification of the dependent variable.

lower French dummy, we prefer to employ the German dummy to correct the other patents.

The application year dummies are not jointly significant in the three equations. Most likely this is because we asked our inventors to evaluate their patents in 2003 by using all the information that they had when they answered. Thus, all our patents are evaluated 6–10 years since their year of application. The value of patents, and of the underlying inventions, may change considerably in their first few years of life either because of rapid technological obsolescence, or because important (positive or negative) information becomes available. But 6–10 years later (or more in terms of priority dates), this value stabilizes. As a result, there may be no difference according to whether the patent was 6 or 10 years old.

Estimated distribution of patent values

In order to better describe our distribution of patent values, we estimated some moments of its hypothetical log-normal distribution. As noted, the literature suggests that the distribution of patent values may not be exactly log-normal because the right tail may be fatter. However, the log-normal distribution assumption may hold for a wide range of patent values before the extreme right tail of the support.

Table 8 uses the estimation results in the first column of Table 7. We consider the fitted values of the regression by using the German constant term for all the observations. This yields an expected value $\mu_i \equiv E[\log(\text{VALUEM})]$ for all our observations. We compute this statistic for our PatVal-EU sample observations (8217 observations) and for all the EPO patents with priority date 1993–1997 that were granted by 2003 (49,941 patents). While some patents with priority years 1993–1997 may be granted later, most of them were granted by 2003, and we can take this set to be a proxy for the universe of patents from which the PatVal-EU sample is drawn.[5] To extend our prediction to this larger set we employ data for all these patents on the four indirect indicators, the technology area dummies, the country of the first inventor, and the application years. We also compute μ, which is the average of all the μ_i across the 8217 PatVal-EU patents or all the 49,941 EPO 1993–1997 patents.

If we assume that the distribution of the error of the first equation in Table 7 is normal, then the distribution of the

Table 8 Estimated moments of the patent value distribution, assuming log-normality (values in 000 2003 euros, corrected by German dummy)

Moment	*Theoretical expression for the log-normal distribution*[a,b]	*Estimated moment*[b,c]
Mean	$exp(\mu+\sigma^2/2)$	*Average of* $exp(\mu_i+\sigma^2/2)$
PatVal-EU sample ($N=8217$)	3138.6	3550.8
All patents ($N=49{,}941$)[d]	3015.6	3422.6
Median	$exp(\mu)$	*Median of* $exp(\mu_I)$
PatVal-EU sample ($N=8217$)	397.4	382.7
All patents ($N=49{,}941$)[d]	381.8	365.3
Mode	$exp(\mu-\sigma^2)$	*Average of* $exp(\mu_i-\sigma^2)$
PatVal-EU sample ($N=8217$)	6.4	7.2
All patents ($N=49{,}941$)[d]	6.1	6.9

[a]The parameter μ is the average of the fitted values of the first equation in Table 7, viz. *E*(log(VALUEM), using the German constant for all the observations. For the PatVal-EU sample the average is computed across the 8217 PatVal-EU observations, and it is $\mu=5.985$. For the full set of patents is predicted from the available regressors for all the 1993–1997 EPO patents, and it is $\mu=5.945$.
[b]The estimated $\sigma=2.033$ is the standard error of the first regression in Table 7.
[c]The parameter μ_i denotes the fitted values of log(VALUEM) for the generic *i*th observation.
[d]All EPO patents with priority year 1993–1997 granted by 2003.

patent values is log-normal. This implies that the mean, median and mode of patent value are the theoretical expressions indicated in Table 8 as a function of the mean and standard deviation of the log-value distribution. Table 8 reports our estimates of these moments as the theoretical expressions computed at the sample μ or as the sample average of the expression evaluated at each observation (μ_i) in the sample (whether PatVal-EU sample or all EPO patents). The results are strikingly similar. Moreover, the estimated moments are similar whether we compute them for the PatVal-EU sample or for all the EPO patents. This suggests that our sampling weight methodology and our empirical analysis more generally adequately represents the universe of EPO patents.

Our estimated mean value of patents is higher than 3 million euros. The median is almost 400 thousand euros, and the mode of the distribution is between 6 and 7 thousand euros. This mirrors several features of the patent value distribution that have emerged from the literature. First, because our measure of patent value includes both the value of the invention and the patent premium, it is naturally higher than estimates based on renewal fees that only cover the patent premium. Second, our estimates are suggestive of the highly skewed nature of the patent value distribution. As noted in the earlier sections, some highly valuable patents can boost the mean value. We claim that, unlike other studies in this field, the PatVal-EU research captures some of the high value patents at the very right tail of the distribution, and this is what makes the mean value of patents quite high. At the same time, the mean of an asymmetric distribution does not coincide with the most 'representative' patent, or the most likely value of the distribution that one can observe. In fact, if one randomly draws patents, it is unlikely that one draws a patent whose value is equal to the mean value. The median, which is smaller given the skewness of the distribution, is not the most likely value to observe either. This is instead the mode, which in our case is quite small, viz. a few thousand euros. In short, our analysis is not in contrast with the common sense that if one draws a patent randomly, it will most likely exhibit a low value. Yet, with asymmetric distributions this by no means implies that the mean-value, or even the median, is small.

Note also that these mean values are smaller than the mean value of the PatVal-EU sample values (10.9 million euros from Table 2), and of the German PatVal-EU average of 5.6 million euros, which is a better comparison given the German correction that we introduced in our estimates. This suggests that our regression approach probably underestimates the mean of the distribution, consistent with the view that in the case of patent values the log-normal distribution may not fully take into account the fatness of the right tail, and that the applied indirect indicators of value are in fact imperfect.

Finally, we computed the estimated means and medians by sector. They are computed as earlier with μ_i now indicating the left-hand side of the equation for the observations regarding the corresponding sector. The median is computed for the observations within the sector. We find that the most valuable patents (median or means) are in the chemical-related industries and in pharmaceuticals, which confirms the finding of previous studies (e.g., Levin *et al.*, 1987).

The relationships between patent value and CITES, REFS, CLAIMS, STATES

Explanatory power of CITES, REFS, CLAIMS, STATES

Table 9 reports the changes in R^2 as we gradually introduce our indicators as covariates in the value equation (first column of Table 7). These are changes in the extent to which the addition of our regressors explain the variance of log(VALUEM).[6] As shown by the table, the addition of CITES improves R^2 by 1.4%, that is, from 6.2% – which is the variance explained only by the country, technology, and application year dummies – to 7.9%. REFS and CLAIMS have a very small additional explanatory power, while STATES has a somewhat larger impact (0.9%).[7]

The overall impact of our four indicators is therefore small (2.7%). However, *F*-tests not reported in Table 9 indicate, that all the impacts, whether individual or combined, are statistically significant. In brief, the impact of our indicators on patent value is systematic and not random, but it is small. As a benchmark, 4.1% of the 6.2% of the variation explained by country, technology, and application year dummies is produced by the country dummies, and 2.3% by the technology dummies. This suggests that controlling for country and technology can be even more important than our indicators.

In this spirit, the last column of Table 9 presents regression results obtained after adding 116 dummies (one omitted) for the three-digit IPC classes of our patents. As noted in Table 1, the 30 dummies for technological fields are built mostly from the patents four-digit IPC classifications. The ISI-INPI-OST concordance maps the IPC classes into technological classes, but it combines patents from various four-digit classes. Hence, the two sets of dummies can be jointly employed as covariates in a regression. The IPC dummies then reflect strict technological characteristics, while the ISI-INPI-OST classes reflect industry (and possibly market) characteristics. The last column of Table 9 shows that the addition of the IPC three-digit dummies augment R^2 by another 2.1%, which is one-third higher than the addition produced by CITES. This strengthens the importance of controlling for technological characteristics in assessing patent value.

Even after including the IPC three-digit dummies we only explain 11.3% of the variance of patent value. Since the state-of-the-art in the economics and management literature is to measure the economic value of patents by means of indicators like the ones used in this paper, and by industry or technology-fixed effects, this is suggestive of the breadth of the unexplained residual. To borrow Moses Abramovitz's famous expression, the 'measure of our ignorance' is still sizable. In this respect, our paper highlights the need for new research in this field, with some decisive theoretical or empirical innovations to produce a quantum leap in our understanding of these matters.[8]

Right-tail of citations *vs* right-tail of patent value

Because of the skewed distribution of both patent value and citations, it is natural to ask whether the right tail of citations is correlated with the right tail of patent values. To perform this experiment, we employ the CITES0-5 dummies listed in Table 1. They correspond to the classes of CITES that include the 1st quartile, the median, the 3rd quartile, the 95th percentile, the 98th percentile, and the 99th percentile of the distribution of citations. Similarly, we employ the three dummies VALUE$\geqslant$5, VALUE$\geqslant$6, and VALUE$\geqslant$7, which take the value 1 if the patent falls in our value classes 1–3M or above, 3–10M, or above, more than 10M, respectively. The first dummy is equal to 1 for 17.5% of our 8217 PatVal-EU patents, while the second and third dummies are equal to 1 for 7.8 and 3.9% of the patents.

Table 10 reports the results of four regressions. The first column is an OLS regression with log(VALUEM) as the dependent variable. The covariates are the logs of REFS, CLAIMS, and STATES. Since we showed that IPC three-digit and ISI-INPI-OST dummies have independent explanatory power, we include both sets of dummies, along with country and application year dummies. Finally, we include the five dummies CITES1–5 (CITES0 omitted) to account for different effects of the CITES-classes particularly at the tail of the citation distribution. The other three columns report the marginal effects of probit regressions for the three top-value dummies, and the same covariates.

Table 9 Testing the impact of CITES, REFS, CLAIMS, and STATES, OLS regressions, dependent variable log(VALUEM)

	Model I	*Model II*	*Model III*	*Model IV*	*Model V*	*Model VI*
CONST	6.614***	6.374***	6.063***	5.678***	5.066	5.823***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes
Technology class dummies	Yes	Yes	Yes	Yes	Yes	Yes
Application year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Log(1+CITES)	—	0.396***	0.385***	0.361***	0.343***	0.356***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(1+REFS)	—	—	0.198***	0.171***	0.158**	0.137**
			(0.001)	(0.006)	(0.010)	(0.030)
Log(CLAIMS)	—	—	—	0.193***	0.171***	0.170***
				(0.000)	(0.000)	(0.000)
Log(STATES)	—	—	—	—	0.372***	0.395***
					(0.000)	(0.000)
IPC 3-Digit Dummies	—	—	—	—	—	Yes
R^2	0.065	0.079	0.080	0.083	0.092	0.113
Change in R^2	—	0.014	0.001	0.003	0.009	0.021
N. Observations	8217	8217	8217	8217	8217	8217

P-values based on robust standard errors in parentheses. *$P<10\%$; **$P<5\%$; ***$P<1\%$. All regressions use sampling weights, and observations are clustered by patent applicants. The table shows that progressive addition of citations and other indicators as regressors improves fit by a few percentage points. Citations improve fit by 1.4%, which is just slightly higher than the contribution of the other three indicators combined. Although statistically significant, the overall impact of the indicators is small. Even after including the IPC 3-digit dummies we can only explain 11.3% of the patent value.

Table 10 Testing the impact of the tail of CITES on the tail of VALUEM

	Dependent variable log(VALUEM) OLS	*Dependent variable VALUE⩾5 Probit (marginal effects)*	*Dependent variable VALUE⩾6 Probit (marginal effects)*	*Dependent variable VALUE⩾7 Probit (marginal effects)*
CONST	5.835***	—	—	—
	(0.000)			
CITES1	0.152**	0.017	0.006	−0.000
	(0.011)	(0.101)	(0.342)	(0.989)
CITES2	0.204***	0.016	0.005	0.002
	(0.005)	(0.213)	(0.560)	(0.670)
CITES3	0.602***	0.076***	0.041***	0.013***
	(0.000)	(0.000)	(0.000)	(0.007)
CITES4	0.732***	0.103***	0.053***	0.019*
	(0.000)	(0.001)	(0.002)	(0.065)
CITES5	1.142***	0.194***	0.064***	0.055***
	(0.000)	(0.000)	(0.006)	(0.001)
Log(1+REFS)	0.136**	0.018	0.009	0.004
	(0.031)	(0.106)	(0.136)	(0.201)
Log(CLAIMS)	0.168***	0.019***	0.005	−0.001
	(0.000)	(0.006)	(0.274)	(0.675)
Log(STATES)	0.395***	0.048***	0.022***	0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.114	—	—	—
N. Observations	8217	8143	8007	7344

P-values based on robust standard errors in parentheses. **P*<10%; ***P*<5%; ****P*<1%. All regressions include country, application year, industry and IPC three-digit dummies, sampling weights, and clustering by patent applicants. In the probits some IPC three-digit dummies perfectly predict the dependent variable, and the corresponding observations are dropped. However, the results shown here are robust to several alternative estimations. The table shows that the right tail of the citation distribution predicts increasingly higher patent values (first column) and that it highly correlated with the probability that the patent falls in the top VALUE classes.

This enables us to estimate the additional probability of being in a top value class if the patent is in a top citation class.

The first column of Table 10 shows that the higher citation classes have an increasing impact on patent value. Figure 4 reports the estimated value of patents at different classes of citations when all else is held constant and the value of a patent in class CITES0 is normalized to 1. Since the dependent variable of this regression is in logarithms, the values in the *y*-axis of Figure 4 are the exponentials of the estimated impacts of the citation classes. As the figure clearly shows, patent value increases exponentially as we move towards the right tail of the citation distribution. This is also implicit in the fact that the estimated impacts of the citation classes in the first column of Table 10 tend to increase as we move towards the higher citation dummies. All this suggests that the right tail of the skewed distribution of citations predicts increasingly distant patent values. In other words, the long right tail of citations predicts a long right tail of patent values.

The other three columns of Table 10 offer a similar picture. The top citation classes and particularly the very top percentiles of citations (CITES4 and CITES5) are good predictors of the top patent value classes. Since the estimated parameters in the table are marginal effects, we can retrieve the probabilities of being in a top value class for different citation classes. These probabilities are reported in Table 11. The benchmark is the baseline probability for the hypothetical case in which citations have no impact, which is clearly 17.5, 7.8 and 3.9% for our three top value dummies. As the table shows, the probability that a patent with citations below the median (CITES0) falls in the top 17.5% patent value classes is 15.4%, that is, just below the random event. This probability increases by more than 100% in the case of a patent in the top 99% of the citation distribution (CITES5), that is, to 34.8%, which is twice as much as the random event. The probability is 23.0 and 25.7% in the cases of CITES3 and CITES4. Interestingly, the probability that a CITES5 patent falls in the top 3.9% is 9.1%, again well above the random event.

While these results clearly suggest that the right tail of the citation distribution is correlated with the right tail of patent values, they also confirm that the noise in the analysis of patent values is significant. While 9.1% of the CITES5 (top 99%) patents fall in the top 3.9% of patent values, and 34.8% fall within the top 17.5%, the remaining 65.2% CITES5-patents, that is, the large majority, have lower values. Likewise, from Table 11, 15.4, 6.9, and 3.6% of the patents with no citations fall in the top 17.5, 7.8, and 3.9% patents ranked in terms of value. In other words, the probability that a patent with no citations falls in the top-value classes is not very far from the random event. In sum, while citations are correlated with the right tail of patent values, there are many (very) valuable patents with no citations, and low-value patents with lots of citations. Our ignorance about the determinants of patent value is not just about the average of the distribution, as noted earlier, but also about its right tail.

The correlation between high citations and high patent values that we observe is consistent with the findings by

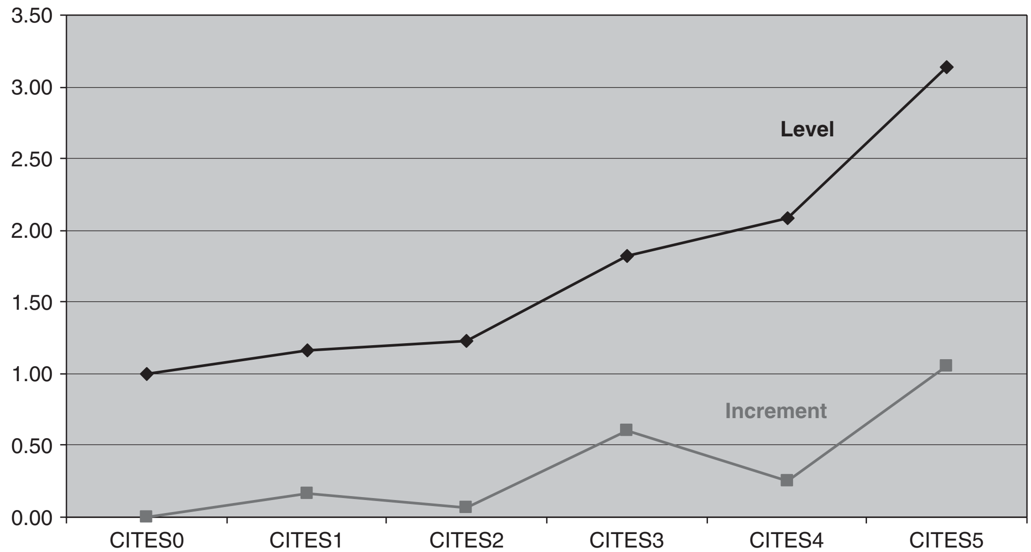

Figure 4 Estimated impacts of the citation dummies (CITES0–5) on patent value (first column of Table 10), impact of CITES0 normalized to 1. Impact of CITES1–5 computed as the exponential of the corresponding estimated parameter in the first column of Table 10. Increment is the change *vis-à-vis* the previous class of CITES. The figure shows that patent values grow exponentially as we move to higher citation classes. Also, increment increases.

Table 11 Estimated probabilities of patents being in the top value classes conditional upon citation class

CITES classes (%tiles)	*Prob(VALUE ⩾ 5|CITES) (top 17.5% patent values)*	*Prob(VALUE ⩾ 6|CITES) (top 7.8% patent values)*	*Prob(VALUE ⩾ 7|CITES) (top 3.9% patent values)*
CITES0 (1–45)	15.4%	6.9	3.6
CITES1 (46–70)	17.1	7.5	3.6
CITES2 (71–83)	17.0	7.4	3.8
CITES3 (84–96)	23.0	11.0	4.9
CITES4 (96–98)	25.7	12.2	5.5
CITES5 (⩾99)	34.8	13.3	9.1

Prob(VALUE ⩾ X|CITES0) $= P^* - \sum_{i=1}^{5} b_i w_i$, where $X = 5$–7; $P^* = 17.5$, 7.8 or 3.9%; b_i is the estimated marginal effect of the CITES classes $i = 1$–5 in the column of Table 10 corresponding to $X = 5$–7; w_i is the share of the CITES class in the 49,941 EPO patent distribution (respectively 25.1, 12.8, 12.3, 2.6, 1.3% for CITES1–5). The probabilities conditional upon CITES1–5 are then computed by adding the corresponding estimated marginal effect. The table confirms that patents in the right tail of the citation distribution are also more likely to be in the top value classes. Yet, quite a few patents with high citations have low value and vice versa. For instance, of the patents in CITES5, 34.8, 13.3 and 9.1% fall in the top 17.5, 7.8 and 3.9% patents ranked by VALUE, which is a much higher probability than the random patent. Yet, this also means that for the vast majority of CITES5 patents (65.2%) VALUE < 5.

Hall *et al.* (2005). They also find that, other things being equal, firms with citations per patent around the median of the distribution of this variable across firms have roughly the same market value as firms below the median. The impact is progressively higher as we move beyond the median, and it becomes quite high as we move close to the right tail of the distribution. Thus, our analysis mirrors, at the level of the individual patents, what they observe at the level of firms.

Conclusions

We employed an unusually comprehensive data set of inventor responses to questions about the economic value of patents. Our estimates are in line with the existing research suggesting that the distribution of patent values is highly skewed. Compared to previous research, our contribution is that we obtain monetary estimates of the moments of the distribution. We find that the mean value of EPO-granted patents is about 3 million euros, the median is about one-tenth of it and the mode is just a few thousand euros. We also find that patent value is significantly correlated with indirect indicators of innovation quality. Moreover, the right tail of the citation distribution is correlated with the right tail of our measure of patent value.

We also find that the overall impact of these indicators is small. More generally, even if our analysis employs detailed fixed effects for countries, industries, and technologies, we explain only 11.3% of the variance of patent value. At the same time, there are many highly cited patents of low economic value, and many patents with no citations and high value. The former may be patents of high social but low private economic value, while the latter may feature high private but low social value. However, at this stage statements like this are confined to the realm of speculation, or they are a plea for future work. The ‘measure of our ignorance' in this field is still too high. This paper finds that new and better explorations of the determinants of the economic value of patents are an important and largely underdeveloped area for future research.

Notes

1 We benefited from comments and suggestions from Ashish Arora, James E. Bessen, Iain Cockburn, Wesley Cohen, Liran Einav, Dominique Guellec, Bronwyn Hall, Thomas Hellman, Jacques Mairesse, Andrea Ordanini, Susanne Prantl, Scott Stern, Manuel Trajtenberg, and Rosemarie Ziedonis. We also thank seminar participants at the UC Berkeley Conference on the Economics of Competition and Innovation, the World Bank-CREI Conference on R&D and Innovation, Universitat Pompeu Fabra, Barcelona; the NBER Summer; the ZEW Conference on Economics of Innovation and Patenting as well as conference and workshop audiences at Bocconi University, Milan; the Tanaka Business School, Imperial College, London, and the SFB/TR15 Conference at Frauenchiemsee. Last but not least we received excellent comments from the Editor of this journal and two anonymous referees. All responsibilities remain ours. We thank the European Commission, Contract N. HPV2-CT-2001-00013, for supporting the creation of the PatVal-EU data set. A.G. also acknowledges financial support from the Italian Ministry of University Research and Bocconi University. D.H. acknowledges financial support from the Deutsche Forschungsgemeinschaft through its SFB/TR 15 program (Project C2).

2 Data for Denmark and Hungary (495 and 38 patents respectively) were obtained in a second round of the survey, and they are neither discussed nor used in Giuri *et al.* (2007). However, the way the Danish and Hungarian surveys were conducted, the sampling criteria, and all the other relevant issues are fully consistent with the details provided in that paper.

3 As discussed in Giuri *et al.* (2007) the PatVal-EU survey deliberately oversampled important patents (patents with at least one forward citation and that were opposed) to increase the sample of valuable patents. The patent value distribution discussed in this section may then overestimate population statistics for this reason as well. However, Giuri *et al.* (2007) show that the impact of this oversampling is small. Moreover, our regression-based estimations in the fifth section use sampling weights to control for potential sample selection effects.

4 The sampling weights are constructed by taking into account PatVal-EU oversampling of important patents and other potential sources of PatVal-EU sample selection (e.g. non-response biases). Details on how the sampling weights are constructed are available upon request.

5 The censoring of the granted patents is not necessarily correlated with their value. More valuable patents may be granted more quickly because they are clearly useful and novel, or more slowly because they encounter more opposition, and are assessed more carefully. We can then fairly assume that the censoring is not a major source of bias, and our extended sample is a good proxy of the universe of patents with priority years 1993–1997.

6 In our analysis the changes in R^2 are very similar to the changes in the adjusted-R^2 because we have a large number of observations.

7 Of course, the order with which we add the repressors affects the change in R^2 produced by each of them. However, we experimented with different orders and found that CITES and STATES tend to have higher impact, than REFS and CLAIRS.

8 Our survey-based measure may add volatility because of subjective estimations of patent value by the respondent. Even though more objective market-based measures may be less fickle, it is unlikely that all our residual is explained by subjective assessments. Interviews with licensing managers that we conducted on several occasions suggest that the market price of technologies also tends to be quite volatile, even for very similar technologies.

References

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

Almeida, Paul and Bruce Kogut, 1999, "Localization of knowledge and the mobility of engineers in regional networks". *Management Science*, **45**: 905–917.

Arora, Ashish, Marco Ceccagnoli and Wesley Cohen, 2008, "R&D and the patent premium". *International Journal of Industrial Organization*, forthcoming.

Bessen, James, 2006, "The value of U.S. patents by owner and patent characteristics". Working Paper Series on Law and Economics, No. 06-46, Boston University School of Law.

Bessen, James, 2007, "Estimates of firms' patent rents from firm market value". Draft.

Gittelman, Michelle and Bruce Kogut, 2003, "Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns". *Management Science*, **49**: 366–382.

Giuri, Paola, Myriam Mariani, Stefano Brusoni, Gustavo Crespi, Dominique Francoz, Alfonso Gambardella, Walter Garcia-Fontes, Aldo Geuna, Raul Gonzales, Dietmar Harhoff, Karin Hoisl, Christian Lebas, Alessandra Luzzi, Laura Magazzini, Lionel Nesta, Onder Nomaler, Neus Palomeras, Pari Patel, Marzia Romanelli and Bart Verspagen, 2007, "Inventors and invention processes in Europe: Results from the PatVal-EU survey". *Research Policy*, **36**: 1107–1127.

Griliches, Zvi, 1981, "Market value, R&D and patents". *Economics Letters*, **7**: 183–187.

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

Harhoff, Dietmar and Karin Hoisl, 2007, "Institutionalized incentives for ingenuity. Patent value and the German Employees' Invention Act". *Research Policy*, **36**: 1143–1162.

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.

Harhoff, Dietmar and Markus Reitzig, 2004, "Determinants of opposition against EPO patent grants – The case of biotechnology and pharmaceuticals". *International Journal of Industrial Organization*, **22**: 443–480.

Harhoff, Dietmar, Frederic Scherer and Katrin Vopel, 2003a, "Exploring the tail of the patent value distribution". In O. Granstrand (Ed.) *Economics, law and intellectual property: Seeking strategies for research and teaching in a developing field*. Dordrecht, Netherlands: Kluwer Academic Publisher, pp: 279–309.

Harhoff, Dietmar, Frederic Scherer and Katrin Vopel, 2003b, "Citations, family size, opposition and the value of patent rights – evidence from Germany". *Research Policy*, **32**: 1343–1363.

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

Levin, Richard, Alvin Klevorick, Richard Nelson and Sidney Winter, 1987, "Appropriating the returns from industrial R&D". *Brookings Papers on Economic Activity*, **14**: 551–561.

Pakes, Ariel, 1985, "On patents, R&D, and the stock market rate of return". *The Journal of Political Economy*, **93**: 390–409.

Pakes, Ariel, 1986, "Patents as options: Some estimates of the value of holding European patent stocks". *Econometrica*, **54**: 755–784.

Schankerman, Mark and Ariel Pakes, 1986, "Estimates of the value of patent rights in European countries during the Post-1950 period". *Economic Journal*, **97**: 1–25.

Scherer, Frederic and Dietmar Harhoff, 2000, "Policy implications for a world with skew-distributed returns to innovation". *Research Policy*, **29**: 559–566.

Scherer, Frederic, Dietmar Harhoff and Joerg Kukies, 2000, "Uncertainty and the size distribution of rewards from technological innovation". *Journal of Evolutionary Economics*, **10**: 175–2000.

Schmoch, U. and N. Kirsch, (1993), "Analysis of International Patent Flows". Karlsruhe Fraunhofer Institut für Innovations Forschung and System technik Karlsruhe.

Serrano, Carlos, 2005, "The market for intellectual property: Evidence from the transfer of patents". Unpublished manuscript. University of Toronto, http://www.economics.utoronto.c2/serrano/papers/measuringtransfers_paper.pdf.

Silverberg, Gerald and Bart Verspagen, 2007, "The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance". *Journal of Econometrics*, **139**: 318–339.

Singh, Jasit, 2008, "Distributed R&D, cross-regional knowledge integration and quality of innovative output". *Research Policy*, **37**: 77–96.

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