World Patent Information 38 (2014) 19–26

Contents lists available at ScienceDirect

World Patent Information

journal homepage: www.elsevier.com/locate/worpatin

Composite value index of patent indicators: Factor analysis combining bibliographic and survey datasets

Grid Thoma*

University of Camerino, School of Sciences and Technology, via del Bastione 3, 62032 Camerino, MC, Italy

ARTICLE INFO

Article history:
Available online 11 June 2014

Keywords:
Patent indicators
Patent market value
Composite value index
Renewals
Opposition

ABSTRACT

The article relates to a patent value composite index that combines twenty different patent indicators according to several dimensions: patent breadth and technology potential, prior art and background of the invention, and filing and procedural aspects of a patent. A novel selection approach of patent indicators and their validation with market value of patents is advanced, whereas the computation of the composite value index is based on the factor analysis methodology firstly proposed in the literature by Lanjouw and Schankerman (2004).

This study reports several new findings. Firstly, three common factors are obtained in correspondence to the respective three dimensions of the patent indicators. Secondly, the proposed composite value index can effectively summarize the information conveyed by every single indicator, because the reduction of the goodness-of-fit of the market value model is very limited as compared to the case of the indicators taken separately. Thirdly, a robustness analysis of the composite value index was conducted, relying on post-grant information, such as opposition and renewal decisions, and the results are consistent with the market value model with the composite value index solely considered. More generally, these findings contribute to the research agenda on proposing novel timely indicators of innovation activities.

1. Introduction

Recent contributions on economic valuation of patent assets have advanced our understanding about the intrinsic value of a technology [1,2]. A first group of studies have estimated the value of patent rights relying on renewal decisions of a patent [3,4]. Other contributions have focused on the market value return of R&D investments and patenting for publicly listed firms [5–7]. Lastly, patent assets have been analyzed in connection to the external financing of start-ups [8]. Nevertheless, the current debate is far from reaching definite answers on how to gauge the heterogeneity of patent value, and what are the determinants of the market value of the R&D output [9].

This paper aims to fill this gap by proposing a new composite index of patent value which aggregates several patent indicators that are typically used in the literature [1] with respect to patent breadth and technology potential, prior art and background of the invention, and filing and procedural aspects of the patent. Overall twenty different indicators are validated with market value information of patents obtained from a survey [10]. Subsequently, I aggregate the selected indicators into a single composite value index following the methodology proposed by Lanjouw and Schankerman [11]. I show that the resulting composite value index summarizes effectively the variability of the indicators taken separately, because the reduction in the explanatory power of the estimated market value model is only about ten percent, which is quite limited considering that I have combined eight distinct indicators.

Furthermore, relying on opposition and renewal decisions I conduct a robustness analysis of the composite value index and the survey's market value dataset. I opted to keep oppositions and renewals apart from other patent indicators, since they are known only after the grant of a patent. I do find that both of them are positively correlated with the market value of patents, confirming the validity of the survey dataset. The inclusion of the composite value index doubles the explanatory power of the estimated model. Conversely, when the composite value index is solely considered, we have that adding the oppositions and renewals increases the goodness of fit of the model of only 14–16 percent. Put differently, not considering post-grant information implies a loss in the

* Tel.: +39 (0) 737 40 25 74.
E-mail address: grid.thoma@unicam.it.

http://dx.doi.org/10.1016/j.wpi.2014.05.005

explanatory power of the predictive model comparable to the situation of all patent indicators taken separately.

2. Background

Previous literature has provided estimates of the value of patent rights, relying in large extent on the patent fees renewal approach [12]: to keep a patent alive a patenter needs to pay fees, and typically not all patents are kept in force until the end of their statutory life. Then, the fee costs structure could be considered as lower bound revenue that the patent can ensure to the owner. Schankerman [3] provides mean and median value estimates of European patents in five industries, analyzing aggregate data of renewal decisions over three decades period. With respect to the US patents, Bessen [4] advances a simulation approach to estimate additional moments of patent value distribution. Other studies measure the return of the decision to patent an invention using survey information [13]: they find that the so called “patent premium” is on average negative, although in some key knowledge intensive industries the value of a patented invention is almost double than the counterfactual situation.

The asset value of patents has been analyzed through the return of the R&D investment and patenting on output variables such as, Tobin's Q or productivity. Many of the studies of this body of literature have regarded the US context [5,11], demonstrating that there is a significant impact on firm's market value of R&D and patents. More recently, some other works have found similar results for the European listed firms [6,7,14], but positing the caveat that European patents are more valuable on average than the US patent counts [14,15]. Greenlagh and Rogers [16] show for the first time that not only patent counts but also trademarks have a significant return on the market value of a firm. To account for the strong skewness of the value of patents, indirect indicators have been proposed such as patent citations, breadth of patent protection in multiple jurisdictions, claims and others. These indicators correlate significantly with firm's market value in addition to R&D investment and patent counts. Lanjouw and Schankerman [11] device a composite value index of patents to account for the multiplicity of patent indicators: this methodology has the advantage of not demanding expert information for patent valuation, and it can solve an aggregation problem of the existing patent indicators.

Another stream of literature has originated from entrepreneurial finance – for a survey see Hall and Harhoff [8] – which stresses the role of patents as quality signals between the inventor and the potential financier. Hsu and Ziedonis [17] show that patents improve the terms by which new firms access venture capital. In particular, they document that the larger the patent portfolio of startups, the bigger the money evaluation by a venture capitalist, and that the effect is even more pronounced for younger and inexperienced firms. In the same vein, Häussler et al. [18] demonstrate similar findings and claim that the granting decision by the patent office does not trigger additional financial evaluation for a venture capitalist, because this event is fully anticipated thanks to innovation indicators revealed in the patent application (e.g. such as patent citations). Other recent papers [19,20] estimate hazard or binary choice models to investigate the relation between patents and company survival, whereas Debb [21] argues that patents are considered credible signals by external investors, and hence they directly affect the financial performance of a firm. Pederzoli et al. [22] propose a model for the estimation of the probability of default including a full-fledged set of financial and patent variables: they show that patent indicators can significantly improve the accuracy of the predictive default model.

Nevertheless, despite the growing number of contributions on patent valuation, the direct analysis of the determinants of the market value of a patent asset is far from reaching full-fledged answers to the debated research agenda. Typically previous studies have relied on survey estimates to gauge market value information: in spite of the high cost of data collection, this methodology can offer more precise estimates of the value of a patent asset as compared to other approaches [2]. Scherer and Harfoff [23] showed that the value of patent assets is extremely skewed with the top decile of the patents holding 84% of the economic value, and they stylized this finding with the log normal distribution. Quite interestingly, the value of patents is found to be one order of magnitude bigger as compared to value estimates of patent rights computed through the renewal approach (see above). Harhoff et al. [24] argued that counts of patent citations can mimic significantly the tail of the patent value distribution. However, the relationship is mediated also by the patent premium: citation counts of patents maintained till to the end of their statutory life have a three times larger return on value than other patents on average. More recently, other studies have argued that citations are characterized by a curvilinear relationship with respect to value, with fewer citations at the tail of value distribution than at the middle [9].

Beyond patent citations, other patent indicators have been claimed to correlate with the value of patent assets. Analyzing a panel of patents invented in Germany of the 1977's cohort and maintained up to the full-term, Harhoff et al. [25] find that successful defense against opposition action is a particularly strong predictor of patent value. They argue that the survival to a two-tier selection process such as grant and opposition is a highly reliable indicator of value. Other variables matter as well including international priorities of patents, backward citations provided during the examination, and – for science based inventions – non patent references. One limitation of this study consists in restrictive assumptions required for extrapolating market value information also for those patents which have not been targeted by the survey.

This methodology has been extended and fully developed by Gambardella et al. [10], who administrated a questionnaire to inventors at a pan European level and they overcame any potential restrictive assumption by scrutinizing a representative sample of the population of patents before the end of their statutory life (see Section 3.4 for further details on this survey). Gambardella et al. argue that patent indicators can predict ex-ante the asset value of a patent, including patent references (backward and forward), patent families as measured by the number of jurisdictions in which the protection is sought, and number of claims, although the unexplained variability of the patent value distribution is considerable. Furthermore, they find that patent citations that could mimic the skewness of the value distribution. These results have been confirmed by Fischer and Leidinger [26] who have built up an econometric model relying on actual market value information originating from Ocean Tomo patent auctions.

It is noteworthy that these studies are largely silent about the impact of innovator's strategic behavior on patent value. More in general, it can be argued that the innovator can put in place strategies in order to ameliorate the conditions that directly affect the value of an invention. For example, Gambardella et al. [27] question to what extent patent value depends on the effort dedicated to the underlying R&D project – as measured by the workload of inventors – with respect to the innovative breadth of the project, which is given by the number of patent fillings related to the same invention (being part of the same patent family). They find that the second effect is larger than the former and it is mitigated by the technological field of the invention (science based vis-à-vis traditional sectors), and the mode of the invention process, e.g. decentralized technological problem solving as compared to non collaborative R&D projects.

3. Data and variables

This section describes data sources and variables that are then included in the multivariate econometric analysis. The data sources have originated from different repositories: i) patent indicators from the EPO Worldwide Patent Database; ii) opposition decisions from the EPOLINE XML files; iii) procedural information on renewals of patents from the EPO Patent Register (EPR) Database; iv) market value estimated of patents from survey information. I conclude this section by discussing a methodology for detrending patent indicators for time and technology effects.

3.1. Patent indicators

The main source of patent indicators has been the EPO Worldwide Patent Database [28] and the related Patent Register Data regarding procedural information (see also section 3.3). From PatStat, bibliometric information has been extracted regarding claims, references, patent classifications, inventors, and renewal decisions. Due to data availability, in the econometric analysis the sample is limited to patents originating from the European Patent Office (EPO). In fact, for this dataset I have complete procedural information on applications, grants, oppositions and renewals. Limiting the sample to one single patent office allows more homogeneity and precision in the definitions and computations of the variables. In the Appendix I report the patent indicators of the econometric multivariate analysis.

3.2. Opposition

Opposed patents are more likely to create potential economic losses for opponents, thus have bigger value [25]. Indeed, the average cost of an opposition action has been estimated around €25 000, and the opposition action can account for the potential litigation risk connected to that invention [29]. To this end I have built a binary variable whether a patent has been opposed using information extracted from the weekly XML files available from EPOLINE [30]. These files have been parsed and structured in a SQL relational database, and then linked with the patent indicators dataset through the publication numbers.

3.3. Renewals

The patent renewals consist in the payment of renewal fees every fixed number of periods, which – in the case of European patents are annual periods. As not all patents – are maintained until the end of their statutory life, then the fee costs structure could be considered as lower bound revenue that the patent can ensure to its owner. In this study I make use of the measure proposed by Van Pottelsberghe de la Potterie and Van Zeebroeck [31], which reads as the following:

$$R = \frac{\sum_{t=1}^{T}\sum_{d=1}^{D} G_p(d,t)}{D \times T} \quad (1)$$

where D is the number of designated countries in which the patent p is valid in time t, and T coincides with the years of maintenance of a patent, i.e. the statutory life. In addition $G(.)$ is assumed to be dependent on the GDP of the country where the protection is sought and renewed, which can approximate for the potential market size of an invention. In the current analysis, information on renewal decisions is drawn from the EPO Patent Register Database [28] whereas the national GDP information originates from the Penn World Table [32].

3.4. Survey data

For the validation of the patent indicators, I consider the market value information of patents originating from the survey discussed by Gambardella et al. [10]. This dataset includes value estimates for about 8277 European granted patents through a questionnaire directly administrated to the population inventors resident in the eight largest countries Europe. Gambardella et al. [10] obtain estimates of patent value by administrating the following question to the first listed inventor in a patent: "*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?*" (page 70). Ten intervals of patent values in current Euros are provided: less than € 30K; 30–100K; 100–300K; 300K–1M; 1–3M; 3–10M; 10–30M; 30–100M; 100–300M; more than 300M. It is noteworthy, that the survey is statistically representative at the country and technology level for the cohorts of patents 1993–1997. Patents have been oversampled among those that have been at least cited once or opposed. In the regressions we used weights to account for oversampling of most valuable patents and for potential respondent bias at the country level.

3.5. Detrending indicators for time and technology effects

Before taking the logarithms of the value estimates and patent indicators, I have detrended them by the geometric mean in order to remove technology and time effects. In particular, the resulting detrended log indicator reads as follows:

$$I_{pct}^{D} = Log\left(1 + \frac{I_{pct}}{e^{\frac{\sum_{i=1}^{n}\log\left(1+I_{ict}\right)}{n}}}\right) \quad (2)$$

where I_{pct} is the patent indicator of the focal patent p in technology class c and year t, I_{ict} is the indicator of a given patent i in c and t, and n is the number of all patents in same class and year.

The direct computation of (2) however is limited by the fact that the identification of a unique technological class of a patent is not always viable. Whereas for the US patent documents examiners typically define a single international patent class, in other patent offices – such as the EPO – the international patent classes can be multiple and their order of listing in a patent document does not necessarily follow a technological relevance criteria [28]. In addition, the computation of patent indicators at the family level complicates further the problem, because within a patent family the inventions can be classified in multiple patent classes within and between the offices. To this end I computed (2) for all the m classes listed in a patent family and then I took the averages across those classes, having the resulting transformation:

$$I_{pct}^{D} = \frac{\sum_{j=1}^{m} Log\left(1 + \frac{I_{pct}}{e^{\frac{\sum_{i=1}^{n}\log\left(1+I_{ijct}\right)}{n}}}\right)}{m} \quad (3)$$

The transformation (3) has been implemented at several levels of technological disaggregation. In particular, I have followed a computation cascade from level one to level four of the international patent classification. I think that this procedure allows to smooth more effectively technological and time trends, which have been shown to account for a significant share of the variability of patent value [10].

4. Composite value index of patent indicators

I develop a novel patent composite value index relying on several patent indicators, typically used in the literature [1]. In particular I consider three macro groups of uni-dimensional patent indicators: patent breadth and technology potential; prior art and patent background; and filing and procedural aspects of a patent (see the Appendix for more details). Then, I aggregate these variables into a single composite indicator following the factor analysis methodology proposed by Lanjouw and Schankerman [11] and Hall et al. [14], who showed that a similarly built index is positively and significantly associated with the market value of publicly quoted companies. I build upon these stream of studies and I advance improvements regarding the selection and validation (STAGE I) of patent indicators to be used in the computation of the composite value index (STAGE II).

4.1. Selection and validation of patent indicators

The selection and validation stage of the patent indicators [33] has consisted in comparing them with survey based value estimates with the mean of a stepwise regression. As dependent variable I used the log of the mid-point of each interval of the patent value from the survey, that is in current Euros: 15K, 65K, 200K, 650K, 2M, 6.5M, 20M, 65M, 200M, and 650M for the last interval. The log transformation allows to regress the impact of patent indicators on the value estimates by least squares: as robustness check I consider also the log–log transformation. It is noteworthy that the log transformation of the mid-point of each interval provides also a more conservative estimate of the value of patents against various factors: miss-classification of a patent in a given value interval, extreme tail of the value distributions, and presence of outliers and influential observations.

The stepwise regression's results are reported in Table 1 (Model 1–2). Only the patent indicators which are statistically significant at five percent level are shown. All regressions include also two dummies for no backward and forward citations.

We can notice that the selected indicators are eight, and the results are confirmed even when we consider the log–log specification (Model 2). In particular, they are:

i) Patent family given by the number of patents that share the same priority;
ii) Patent family weighted by the GDP of the country where the protection is sought;
iii) Number of claims in the frontpage of a patent;
iv) Number of forward citations received within five years from priority year;
v) Number of backward citations, excluding non patent references;
vi) Number of XY backward citations;
vii) PCT route dummy, indicating whether the patent originates from an international filling;
viii) Supplementary search report dummy.

As in Gambardella et al. [10] I find that patent family sub i) is the main positive predictor of patent value followed more distantly by patent citation variables. The family size indicator weighted by the extent of the national market where protection is sought adds further explanatory power to the model with an elasticity about half of that of the unweighted indicator. This is not a small amount and confirms that the market size of the legislation where protection is sought can be considered a direct demand-size measure of the future profitability of a given invention [25].

The PCT route is positively correlated with patent value, which points to the fact that patent owners typically seek eventual global protection with an international application when they think ex-ante that the focal invention is more valuable. While the decision to start the granting procedure of the international application into the so-called regional phase can be postponed for twenty months or more from filing, the fact that the PCT route is the most expensive type of filing signals the larger economic potential of an invention, and on average it associated with bigger success regarding the granting decision.

Patent citations affect positively patent value, although this effect is more limited than patent family indicators. The forward and backward citations have an overall elasticity of about one-third of the two indicators of patent family taken together. This evidence

Table 1
Log linear regression of patent value with clustered standard errors.

(8277 EPO patents from the survey dataset)						
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
	log(value)	log–log(value)	log(value)	log–log(value)	log(value)	log–log(value)
Costant	0.169*** (0.01)	0.179** (0.005)	0.229*** (0.003)	0.218*** (0.002)	0.245*** (0.006)	0.219*** (0.003)
No forward citations	0.000 (0.003)	−0.002 (0.002)	−0.004** (0.001)	−0.005*** (0.001)	−0.001 (0.003)	−0.002 (0.001)
No backward citations	0.010 (0.007)	0.006** (0.003)	−0.005** (0.003)	0.004* (0.002)	0.012** (0.006)	0.007** (0.003)
Patent family weighted	0.024*** (0.011)	0.013*** (0.006)				
Patent family	0.042*** (0.007)	0.021*** (0.004)				
Forward citations	0.011*** (0.003)	0.005*** (0.001)				
Claims	0.012*** (0.004)	0.007*** (0.002)				
Backward citations	0.005** (0.003)	0.003** (0.001)				
XY backward citations	0.005** (0.002)	0.002** (0.001)				
D (SSR)	−0.070*** (0.032)	−0.036*** (0.014)				
D (PCT route)	0.013*** (0.005)	0.008*** (0.003)				
Composite index (computed in all EPO dataset)					0.028*** (0.002)	0.014*** (0.001)
Composite index (computed in the survey dataset)			0.007*** (0.001)	0.009*** (0.001)		
Adjusted R2	0.052	0.050	0.032	0.031	0.046	0.044
R2 relative reduction			39.4%	37.5%	11.3%	11.8%

Notes:
1) *** indicates statistically significant at 1% level, **5% level, and * 10% level.
2) Standard errors clustered at the level of the patenter.
3) Regressions include statistical weights for respondent bias at the country and oversampling most valuable patents.

Table 2
Descriptive statistics and correlation table of the patent indicators. (EPO patents with priority years 1978–2007 – 2008 134 patent applications).

		Mean	S.D.		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	Patent family weighted	0.527	0.154	(1)	1.000							
(2)	Patent family	0.703	0.337	(2)	0.629	1.000						
(3)	Forward citations	0.565	0.566	(3)	0.438	0.383	1.000					
(4)	Claims	0.694	0.420	(4)	0.130	0.165	0.180	1.000				
(5)	Backward Citations	0.702	0.490	(5)	0.514	0.489	0.552	0.198	1.000			
(6)	XY Backward Citations	0.500	0.456	(6)	0.018	0.047	0.080	0.124	0.166	1.000		
(7)	D (PCT route)	0.414	0.493	(7)	0.163	0.233	0.037	0.071	0.062	0.091	1.000	
(8)	D (SSR)	0.109	0.312	(8)	0.077	0.073	0.079	0.064	0.123	0.195	0.400	1.000

Notes: All coefficients are statistical significant at 5% level. The continuous variables are log transformed.

confirms previous studies who have argued on the noisiness of citations in measuring patent value (see Ref. [34] for a fuller discussion). It is noteworthy that supplementary search report variable - a measure of prior art - is negatively related to patent value; in absolute terms its effect is even larger than forward cites and it is comparable to the other two measures of patent family taken together. The latter result is quite striking and at the best of my knowledge this paper is among the first studies that has demonstrated the relevant role of the supplementary search report to explain patent value [35].

In a non reported regression I analyze the backward citations by different categories, that is the inventor added citations at the moment of filling, XY citations added by the examiner, and other examiner added citations. While the impact of two count variables of examiner citations are positive and highly significant, the number of inventor citations is not significantly correlated with patent value, but there is evidence at ten percent level of significance that patents that include at least one inventor added citation are more valuable. These findings do not contradict the assumption that inventions with a bigger number of inventor citations are more connected to the local dimension of the innovation relatively to the geographic and technological distance, and therefore those inventions could be more derivative in nature and less valuable on average [36]. On the other hand, a bigger number of examiner added citations characterize inventions with larger technological breadth, and typically these inventions are better drafted by the applicant and their law firm representatives, i.e. they already provide prior art at the moment of the filing.[1]

The overall goodness of fit of our model is limited as reflected by the R squared ratio, which is only of about 5.2% not very high but in line with the model proposed by Gambardella et al. [10] who have found that four indicators – forward cites, backward cites, claims and family size – account for about 2.7% of total variability of patent value. The present analysis improves along several dimensions: i) include additional indicators such as weighted family size, XY backward cites, PCT route and supplementary search report; ii) detrend the patent indicators along time effects and technology class by the geometric mean; iii) take into the account the overall technology classes in a patent and not only the primary listed class. As in Gambardella et al. [10], I find that the goodness of fit of the log specification is higher than the log–log transformation, and the difference is of a similar order of magnitude (about 0.2 percentage points).

4.2. Aggregation of the composite value index with factor analysis

The construction of the composite value index of patent indicators relies on the factor analysis (STAGE II). In factor models each series of data is decomposed into a common component driven by some common features and an idiosyncratic component, which represents the remaining unexplained variability in the population. To estimate the factors I use principal component analysis, a methodology that seeks a linear combination of variables that maximizes the variance of the sample for the first factor, then the remaining of the variance for the second factor, and so on. Furthermore, I use a multiple indicators model with the unobserved factors and error terms normally distributed [11]. The assumption of normality allows to estimate by maximum likelihood which ensures the uniqueness of the solution of the factor analysis. Estimation of the factors is based on information extrapolated from the covariance matrix of patent indicators (For a fuller discussion see Ref. [37]).

The descriptive statistics and correlation matrix of patent indicators are reported in Table 2. All indicators are correlated at five percent level of statistical significance. We can notice that family indicators are highly correlated with forward and backward cites, and in part also with the PCT route dummy pointing to the fact that owners rely on the World Intellectual Property Office to achieve broader protection. XY Backward citations are correlated with Claims and with the SSR dummy, which in turn is positively correlated with all the other indicators.

The positive correlation of the SSR dummy with other indicators does not necessarily contradicts the results of Table 1, where it is shown that the SSR dummy has a negative impact on the patent value estimates. Indeed, all indicators are affected by the presence of many null values and small number biases. As expected, the SSR dummy is mainly correlated with the PCT route revealing that in large majority of cases the SSR comes in presence of International Search Authorities other than EPO. By including the PCT dummy in the regressions of Table 1, I can identify the net effect of the SSR on the patent value compared to the other indicators.

The results of the factor loadings are reported in Table 3. We can notice the existence of three factors which can be associated to the macro groups of uni-dimensional indicators proposed in Appendix A respectively:

- F1: Breadth and technology potential mainly constituted by patent family and citations indicators;
- F2: Filing and procedural aspects of the patent based on procedural information of the patent application (PCT route and SSR dummies);
- F3: Prior art and background of the invention such as the two measures of backward citations.

4.3. Robustness analysis

Robustness analyses of the composite index with estimates of patent value from the Gambardella's et al. [10] dataset are also reported in Table 1 (see Models 3–6). I find that the composite

[1] The results of these regressions are available by request from the author.

Table 3
Factor loadings of the principal component analysis with maximum likehood. Overall EPO patents with priority years 1978–2007 - 2 008 134 patent applications.

		Factor 1	Factor 2	Factor 3
(1)	Patent family weighted	0.273	−0.048	−0.255
(2)	Patent family	0.329	0.020	−0.464
(3)	Forward citations	0.158	−0.094	0.198
(4)	Claims	0.036	0.014	0.057
(5)	Backward Citations	0.354	−0.159	0.497
(6)	XY Backward Citations	0.022	0.078	0.148
(7)	D (PCT route)	0.061	0.435	−0.122
(8)	D (SSR)	0.050	0.465	0.209
Eigenvalues of variance covariance matrix		2.257	0.836	0.836

value index summarizes effectively the variability of the indicators taken separately (Model 5–6). The reduction in R squared ratio is of order of about ten percent, which is quite limited considering that I have combined eight distinct indicators. Moreover, estimating the composite value index using the overall EPO dataset provides additional informational improvements compared to the case of reducing the estimation of the index with the mean of a smaller scale dataset (Model 3–4).

The point elasticity of the composite value index is about 0.025, which in terms of one standard deviation increase (0.661) generates an impact of about 1.6% on the log value of the average patent, that is 10.9% growth in absolute value. In the same vein, two standard deviations increase generate 23.3% growth in absolute value of the average patent, and three additional standard deviations are associated with 37.3% bigger value.

As discussed in the previous section, an additional robustness check has been that of including the opposition and renewal decisions, which are considered by the literature relevant measures of patent value [25,38,39]. However, they are known only after the grant of a patent, and for this reason I have opted to keep them separate from other patent indicators, which are the base for the composite value index. Including these two variables in the value regression constitutes a robustness check of the survey dataset as well.

Table 4 reports the results of these regression. As we can notice the opposition and renewal variables are positively correlated with the patent value, confirming the validity of the survey information (Model 7 and 8). The inclusion of the composite value index doubles the explanatory power of the estimated model (Model 9 and 10). However, compared to the case of the index solely considered (Model 5 and 6) the goodness of fit of the model increases of only 14-16 percent. Put differently, not considering as value determinants opposition and renewal decisions which can be traced only after the grant of a patent, implies a loss in the explanatory power of model which is comparable to the case of all patent indicators taken separately.

5. Conclusions

The analysis advanced in this paper contributes to the research agenda on proposing novel timely indicators of innovation activities [40]. I devised a composite value index that aggregates information from several patent indicators regarding patent breadth and technology potential, prior art and background of the patent, and filing and procedural aspects of the patent. Overall twenty different indicators typically used in the previous literature [1] are validated with respect to market value of patents obtained from survey information. The aggregation of these variables into a single composite index is done with the mean of factor analysis. The composite value index summarizes effectively the variability of the indicators taken separately: when I regress the composite index against the market value of patents, the relative reduction in goodness of fit of the predictive model is only about ten percent.

These results are robust when post grant information of a patent such as opposition and renewal decisions are taken into the account. First, I found that both oppositions and renewals are positively correlated with the market value of patents originating from the survey dataset adopted in this study. Second, the inclusion of the composite value index doubles the explanatory power of the estimated model. Conversely, when the composite value index is solely considered, adding opposition and renewal decisions increases the goodness of fit of the model by only 14-16 percent.

Future research could advance in several directions. First of all, it is of high interest to elaborate more fine grained measures of the impact of citations and prior art to patent value and the combination of these variable to other indicators. This could include measures of self-citations at the level of the same patent owner, origin of the cited prior art in terms of sectoral activity of patenter, experience with the patent system, availability of complementary assets and others. Secondly, additional robustness analyses are

Table 4
Log linear regression of patent value with clustered standard errors.

(8277 EPO patents from the survey dataset)				
Dependent variable	(7)	(8)	(9)	(10)
	log(value)	log−log(value)	log(value)	log−log(value)
Costant	0.252*** (0.007)	0.222** (0.003)	0.233*** (0.007)	0.213*** (0.003)
No forward citations	−0.013*** (0.003)	−0.008*** (0.001)	−0.002 (0.003)	−0.002 (0.001)
No backward citations	−0.020*** (0.005)	−0.009*** (0.002)	0.010* (0.006)	0.006** (0.003)
D (No renewals index)	−0.002 (0.005)	−0.001 (0.003)	−0.004 (0.005)	−0.003 (0.003)
Renewals index (log)	0.209*** (0.035)		0.118*** (0.037)	
Renewals index (log−log)		0.119*** (0.02)		0.066*** (0.021)
D (Opposition)	0.010*** (0.004)	0.005*** (0.002)	0.008** (0.004)	0.005*** (0.002)
Composite index (computed in all EPO dataset)			0.025*** (0.002)	0.013*** (0.001)
Adjusted R2	0.024	0.025	0.053	0.052
R2 relative change[a]			116.4%	106.4%
R2 relative change w.r.t. (Table 1)[b]			13.8%	16.5%

Notes:
1) *** indicates statistically significant at 1% level, **5% level, and * 10% level.
2) Standard errors clustered at the level of the patenter.
3) Regressions include statistical weights for respondent bias at the country and oversampling of the most valuable patents.
[a] The comparisons regard Model 7 with 9, and Model 8 with 10.
[b] The comparisons regard Model 5 with 9, and Model 6 with 10.

required with respect to patent datasets originating from other legislations beyond the European Patent Convention. A third direction is related to the aggregation methodology of the composite value index which can take into the account also second order conditions and interaction terms of single patent indicators [41].

Editor's note

The following publications in the field of patent value indicators may also be of interest to readers.

- Kapoor, R., Karvonen, M., Kässi, T. Patent value indicators as proxy for commercial value of inventions. International Journal of Intellectual Property Management, 6(3), 2013, 217–232
- Neuhäusler, P., Frietsch, R. Patent families as macro level patent value indicators: Applying weights to account for market differences. Scientometrics, 96(1), 2013, 27–49
- van Zeebroeck, N. The puzzle of patent value indicators. Economics of Innovation and New Technology, 20(1), 2011, 33–62
- Meyer, M.S., Tang, P. Exploring the "value" of academic patents: IP management practices in UK universities and their implications for Third-Stream indicators. Scientometrics, 70(2), 2004, 415–440
- Reitzig M. Improving patent valuations for management purposes - Validating new indicators by analyzing application rationales. Research Policy, 33(6–7), 2004, 939–957.

Acknowledgments

In this research project I have benefited from comments by David Abrams, Jim Bessen, Xiangdong Chen, Iain Cockburn, Alberto Galasso, Alfonso Gambardella, Akira Goto, Stuart Graham, Dominique Guellec, Bronwyn H. Hall, Dietmar Harhoff, Francesco Lissoni, Michael Meurer, Sadao Nagaoka, Salvatore Torrisi, Bart Verspagen, Maria Pluvia Zuniga, and all the participants at the Innovation Seminar at Beihang University in Beijing (March, 2013), Management of Intellectual Property Workshop at Alma Graduate School in Bologna (June, 2013), the ESF Academic Patenting Conference at ParisTech (September, 2013), and the Patent Statistics Conference for Decision Makers Conference in Rio de Janeiro (November, 2013). Also, this article has been developed in part from my presentation on "The Value of Patent and Trademark Pairs" at the November conference.

Appendix. Uni-dimensional patent indicators.

Group I : Breadth and technology potential	
Patent family	Number of patents that share the same INPADOC priority. Economic value is related to the willingness of the owner to seek protection for the same invention across multiple jurisdictions [10]. When the EPO patent documents are not typically published as internal documents of a European national office, family size has been combined with procedural information on the designation decisions for that country.
Patent family weighted by the market size	As the patent family variable where each jurisdiction has been weighted by the GDP of the country where the protection is sought.
Claims	A count variable of the number of claims of the patent at the moment of grant or application.

(*continued*)

Group I : Breadth and technology potential	
Patent classes	Number of full digits technology classes (European Patent Classification) in which the patent was classified by the patent office [42].
Forward citations	The number of forward cites received by the patent or its equivalents during the first five years [14].
Generality	Herfindhal index measuring the dispersion of forward citing patents over their technology classes (4 digits of the International Patent Classification) during the first five years of the patent lifecycle [43].
Group II :Prior art and background of the invention	
Inventors	Number of inventors in a patent [44].
Backward Citations	Number of citations to other patent documents. A bigger number of cites indicates that an invention relies on a broader knowledge base, and hence it is more important [11].
XY Backward Citations	A count variable of cites made to other patents whose claims and/or subject matter overlap completely or partially with at least one claim of the focal patent [45]. It measures the degree of importance of prior art to the focal patent.
Non patent references	Number of citations to the non-patent references prior art, which proxies the closeness to 'science' knowledge [46].
Originality	Herfindhal index measuring the dispersion of backward citing patents over their technology classes (4 digits of the International Patent Classification) [47].
Citation lag	Measures the average age of the backward citations in days [47].
Group III : Filling and procedural aspects of the patent	
Decision reached	A binary variable that takes the value 1 if the patent has been granted, rejected, or withdrawn. It proxies the complexity and uncertainty of the examination process [48].
Grant	A binary variable that takes the value 1 if a patent has been granted by the office.
Decision lag	Number of days between the priority date and the decision date.
Grant lag	Number of days between the priority date and the grant date.
PCT route	A binary variable that signal whether the patent owner has filed an international application via the Patent Cooperation Treaty agreement, which allows to achieve global protection by filing an unique application.
Supplementary search report	A binary variable if the patent has been accompanied with a supplementary search report by the examiner. In the EPO the examiner can optionally choose to elaborate an additional prior art search, when she or he thinks that patent application still lacks relevant prior art in the matter [45].
Divisional application	A binary variable if the patent has at least one divisional application with a common priority patent.

References

[1] Nagaoka S, Motohashi K, Goto A. In: Hall BH, Rosenberg N, editors. Patent statistics as innovation indicators, in the handbook of the economics of innovation. Elsevier Ebook; 2010.
[2] Bessen J, Meurer M. Patent failure: how judges, bureaucrats and lawyers put innovators at risk. Princeton, NJ: Princeton University Press; 2008.
[3] Schankerman M. How valuable is patent protection: estimates by technology fields. Rand J Econ 1998;29(1):77–107.

[4] Bessen J. The value of U.S. patents by owner and patent characteristics. Res Pol 2008;37:932–45.
[5] Hall BH, Jaffe A, Trajtenberg M. Market value and patent citations. Rand J Econ 2005;36:16–38.
[6] Toivanen O, Stoneman P, Bosworth D. Innovation and the market value of UK firms 1989–1995. Oxf Bull Econ Stat 2002;64:39–61.
[7] Hall BH, Oriani R. Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany. Inter J Ind Org 2006;24:971–93.
[8] Hall BH, Harhoff D. Recent research on the economics of patents. A Rev Econ 2012;4:541–65.
[9] Abrams D, Akcigit U, Popadak J. Patent value and citations: creative destruction or strategic disruption?. NBER Working Paper No. 13426. Cambridge (MA): National Bureau of Economic Research, Inc; 2013.
[10] Gambardella A, Harhoff D, Verspagen B. The value of European patents. Eur Manag Rev 2008;5(2):69–84.
[11] Lanjouw JO, Schankerman M. Patent quality and research productivity: measuring innovation with multiple indicators. Econ J 2004;114:441–65.
[12] Pakes A, Schankerman M. The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources. In: Griliches Z, editor. R&D, patents and productivity. Chicago (IL): University of Chicago Press for the NBER; 1984. pp. 73–88.
[13] Arora A, Ceccagnoli M, Cohen W. R&D and the patent premium. Inter J Ind Org 2008;26(5):1153–79.
[14] Hall BH, Thoma G, Torrisi S. The market value of patents and R&D: evidence from European Firms. NBER Working Paper No. 13426. Cambridge (MA): National Bureau of Economic Research, Inc; 2007.
[15] Belenzon S, Patacconi A. Innovation and firm value: an investigation of the changing role of patents 1985–2007. Res Pol 2013;42:1496–510.
[16] Greenhalgh C, Rogers M. The value of innovation: the interaction of competition, R&D and IP. Res Pol 2006;35:562–80.
[17] Hsu D, Ziedonis R. Patents as quality signals for entrepreneurial ventures, Academy of management best paper proceedings, Anaheim Calif; 2008. http://dx.doi.org/0.565/AMBPP.2008.3653924.
[18] Häussler C, Harhoff D, Müller E. To be financed or not: the role of patents for VC financing. London: Center for Economic Policy Research; 2009. discussion papers No. 7115.
[19] Buddelmeyer H, Jensen P, Webster E. Innovation and the determinants of company survival. Oxf Econ Pap 2010;62(2):261–85.
[20] Motohashi K. Innovation and entrepreneurship: a first look at linkage data of Japanese patent and enterprise census. RIETI Discussion Paper Series 11-E-007, Tokyo; 2011.
[21] Debb P. Signaling type and Post-IPO performance. Eur Manag Rev 2013;10:99–116.
[22] Pederzoli C, Thoma G, Torricelli C. Modelling Credit risk for innovative firms: the role of innovation measures. J Finan Serv Res 2013;44(1):111–29.
[23] Scherer FM, Harhoff D. Technology policy for a world of skew-distributed outcomes. Res Pol 2000;29:559–66.
[24] Harhoff D, Narin F, Scherer FM, Vopel K. Citation frequency and the value of patented innovation. Rev Econ Stat 1999;81:511–5.
[25] Harhoff D, Scherer F, Vopel K. Citations, family size, opposition, and the value of patent rights. Res Pol 2003;32:1343–64.
[26] Fischer T, Leidinger J. Testing patent value indicators on directly observed patent value: an empirical analysis of ocean tomo patent auctions. Res Pol 2014;43:519–29.
[27] Gambardella A, Harhoff D, Verspagen B. What are the determinants of the private value of patented inventions?. London: Center for Economic Policy Research; 2012. discussion paper No. 9462.
[28] PatStat. EPO worldwide patent statistical database – available under license from OECD-EPO task force on patent statistics; April 2012.
[29] Graham S, Harhoff D. Can post-grant reviews improve patent system design? A Twin Study of US and European Patents. London: Center for Economic Policy Research; 2006. discussion papers No. 5680.
[30] EPOLINE. EPO raw product. XML Weekly files available online from: Vienna, Austria: European Patent Office; December 2012 www.epoline.org.
[31] Van Pottelsberghe B, Van Zeebroeck N. A brief history of space and time: the scope-year index as a patent value indicator based an families and renewals. Scientometrics 2008;75(2):319–38.
[32] Penn World Table. National accounts database 1950–2012. Available from: https://pwt.sas.upenn.edu/php_site/pwt_index.php, [accessed December, 2012].
[33] Ernst H, Omland N. The patent asset index – a new approach to benchmark patent portfolios. World Pat Inform 2011;33:34–41.
[34] Gittelman M. The value of European patents. Eur Manag Rev 2008;5:85–9.
[35] Thoma G. Economic value and quality of Chinese patenting: an international perspective. Seoul J Econ 2013;26(1):33–72.
[36] Criscuolo P, Verspagen B. Does it matter where patent citations come from? Inventor vs. Examiner citations in European patents. Res Pol 2008;37(10):1892–908.
[37] Jolliffe T. Principal component analysis. 2nd ed. New York: Springer Verlag; 2002.
[38] Van de Kuilen A. Successful European oppositions: analysis for the patent information professional. World Pat Inf 2013;35(2):126–9.
[39] Laub C. Legal validity metrics for patent offices and law firms – recycling of patent court appeal decisions for quality management. World Pat Inform 2014;37:26–32.
[40] OECD. Entrepreneurship at a glance. Paris: OECD; 2013.
[41] Ozer-Balli H, Sorensen B. Interaction effects in econometrics. London: Center for Economic Policy Research; 2010. discussion paper No. 7929.
[42] Lerner J. The importance of patent scope: an empirical analysis. Rand J Econ 1994;25:319–33.
[43] Hall BH. A note on the bias in the Herfindahl based on count data. Rev Econ Ind 2005;110:149–56.
[44] Guellec D, Van Pottelsberghe de la Potterie B. Applications, Grants, and the value of patent. Econ Let 2000;69(1):109–14.
[45] Hall BH, Thoma G, Torrisi S. Financial patenting in Europe. Eur Manag Rev 2009;6(1):45–63.
[46] Meyer M. Does science push technology? Patents citing scientific literature. Res Pol 1999;29:409–34.
[47] Henderson R, Jaffe A, Trajtenberg M. Universities as a source of commercial technology: a detailed analysis of university patenting 1965–1988. Rev Econ Stat 1998;LXXX(1):119–27.
[48] Harhoff D, Wagner S. Modelling the duration of patent examination at the European Patent Office. Manag Sci 2009;55(12):1969–84.

Grid Thoma (Ph.D, Sant'Anna School of Advanced Studies) is Assistant Professor of Economics and Management at the University of Camerino (Macerata, Italy). He has been visiting professor in several institutions such as the School of Management at Boston University, Graduate School of Engineering at University of Tokyo, School of Management at Ludvig-Maximilian Universitat (Munich), and School of Economics and Management at Beihang University (Beijing). His research interests include innovation and finance in knowledge intensive industries. Recently he has coordinated a research project on patent and trademark datasets with significant impact in the academic community. His website is www.gridthoma.info.