Available online at www.sciencedirect.com
SCIENCE DIRECT®

Research Policy 33 (2004) 939–957

research policy

www.elsevier.com/locate/econbase

Improving patent valuations for management purposes —validating new indicators by analyzing application rationales

Markus Reitzig*

Department of Industrial Economics and Strategy, LEFIC Center for Law and Economics, The Copenhagen Business School, Solbjergvej 3, DK-2000 Copenhagen, Denmark

Received 30 June 2003; received in revised form 30 January 2004; accepted 27 February 2004

Abstract

This paper analyzes the validity of so-far untested indicators of patent value to enhance the quality of patent assessments using indicators at the corporate level. The article expands the theory by eliciting patent attorneys' filing rationales to maximize profits from protecting intellectual property, to inspire the computation of new value indicators, including patent full-texts. Then, based on a newly compiled data set consisting of 813 EP patents, the probability of an opposition against a patent is modeled by established and new value indicators. The results show that accelerated examination requests and qualified word counts enhance the quality of existing valuation methods.

Keywords: Patents; Valuation methods; Value indicators; Strategic patenting; Chemical industry

1. Introduction

> One would be hard-pressed to find a major investment bank that employs even one individual with experience in evaluating patent portfolios. [...] as matters stand now, 'due diligence' regarding patent assets is usually more myth than reality. (Rivette and Kline, 2000)

This critique by Harvard Business Review authors Rivette and Kline is harsh. Existing services offered by investment banking houses to value intellectual property (IP) are given little credit; serious doubts are especially uttered concerning the practitioners' expertise and competence in evaluating patent *portfolios*. At the same time, the authors foresee a rising importance of IP assessment methods in corporate strategy. One may agree or disagree with this statement, however, looking at the scientific literature I find it hard to put the blame on the practitioners. As a matter of fact, despite the diversity of articles from Industrial Organization (IO) or legal scholars on value related issues of intellectual property rights, there is a lack of scientific papers that restructure the knowledge on the evaluation of patent rights from a corporate perspective. Building on earlier works by Pakes (1986) and Harhoff et al. (2003) it turns out that valuation approaches using patent indicators seem especially convenient for the assessment of patent portfolios comprising a large number of intellectual property rights. Here, indica-

* Present address: Center for Corporate Change, Australian Graduate School of Management, University of New South Wales, Sydney, NSW 2052, Australia. Tel.: +45-3815-2535; fax: +45-3815-2540.
E-mail address: reitzig@cbs.dk (M. Reitzig).

0048-7333/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2004.02.004

tors drawing from publicly available patent data banks are computed for individual patents that can then be fed into valuation algorithms yielding the patent portfolio value as the cumulative value of the individual patents (see Lanjouw and Schankerman, 1999, for a related yet not identical approach). Those patent indicators can usually be computed at little cost per patent. However, both from a theoretical and applied standpoint, the indicators need to be valid correlates of patent value. Furthermore, the indicators used should be available early in a patent's life to allow for evaluations of young patents that may be particularly interesting for the corporation's future performance.[1] As the prediction quality of the portfolio's value normally increases with the number of valid patent correlates used in the estimation (unless they are collinear), there exists a vital interest in validating as many indicators as possible.

The scientific challenge at this point therefore lies with the validation of further patent indicators that draw from publicly available information and are available early in the lifetime of the patent. The task is especially aggravated by the complex interdependencies between a patent's economic value, the latent determinants of this value, and observable information resulting finally from legal actions that can be used to compute indicators.

This paper addresses this problem in two steps. First, I provide a theoretical framework by laying out the state of the art and then expanding the existing theory of measuring patent value with indicators. Then, empirical results from a large scale study in the chemical industry are presented. In more detail, the remainder of the paper is structured as follows: Part two of the paper addresses theoretical issues. The third section of the paper describes the research design. Here, the hypotheses concerning the correlation between a patent's economic value and the indicators are derived. Part four of the paper presents empirical results. The paper concludes with a summary providing an outlook on future research.

[1] Note that from an *ex-ante* point of view (filing date of the patent) explanatory variables are already valid patent value indicators if they are correlated with the *anticipated* value of the patent. This paper considers indicators as 'valid' if they either correlate with the patent's value from an *ex-post* or from an *ex-ante* perspective.

2. Theoretical framework

The following section is split in three paragraphs. First, it is worth reviewing briefly a definition of patent value that is suited for corporations that find themselves in a competitive environment. Secondly, an overview of the state of the art on value indicators of patents is presented before I move on to a third subsection in which I try to open up the 'black box' of patent attorneys' work. These descriptions inspire the computation of new value indicators that use observable information from a so-far unexploited source of information in particular: patent full-texts.

2.1. A definition of patent value

The value of individual intellectual assets is rarely observable. Thus, to determine the value of an individual patent, inductive approaches must be chosen and a definition for the latent construct 'patent value' is needed. Harhoff et al. (2003) show in a formalized fashion that for a corporation involved in technological competition, the value of a patent is best defined as its asset value. This definition covers the majority of the empirically relevant scenarios. To determine a patent's value, it is therefore necessary to consider its (observable) effects on *prices*, *costs*, and *sold quantities* of patent-protected products by the owner and its simultaneous (unobservable or counterfactual) effects on the proprietor's competitors. As Reitzig (2003a) shows in a survey of the theoretical literature, counterfactual effects should become assessable when quantifying the following patent's latent value determinants: *state of the art* (*of existing technology*), *novelty*, *inventive step*, *breadth*, *difficulty* of *inventing around*, *disclosure*, and *dependence on complementary assets*. Thus, when speaking of indicators of patent value, they can be theoretically valid correlates of a patent's value in two different cases:

1. either they show a direct correlation with observable *prices*, *costs*, or *sold quantities* of the patent protected product, or
2. they operationalize latent determinants of patent value such as *novelty*, *inventive step*, *breadth*, *difficulty* of *inventing around*, *disclosure*, and *dependence on complementary assets*.

Table 1
Established indicators of patent value

Variable	Validity		Availability in time (months after filing date)	Compilation costs
	Theoretical foundation[a]	Empirical evidence as of today[a]		
Patent age	++	−	48+	E
Market value of corporation	++	++	–	M, partially E
Backward citations	+	+/−	18	E
Forward citations	++	++	Ca. 42+	E
Family size	++	+	18+ (preliminary)[b], ca. 42+ (finally)[c]	E
'Scope'	+	−	18+ (preliminary)[b], ca. 42+ (finally)[c]	E
Ownership[d]	+	++	18+ (preliminary)[b], ca. 42+ (finally)[c]	E
Number of claims	++	+/−	18+ (preliminary)[b], ca. 42+ (finally)[c]	E
Patenting strategy	++	+/−	18 respectively 19+ (preliminary)[b], ca. 42+ (finally)[c]	E
Number of applicants	+	+/−	18+ (preliminary)[b], ca. 42+ (finally)[b]	E
Number of trans-boarder research co-operations	+	+/−	18+ (preliminary)[b], ca. 42+ (finally)[c]	E
Key inventors	+	+	18+ (preliminary)[b], ca. 42+ (finally)[c]	E
Legal disputes (opposition in particular)	++	+/−	ca. 42+ (preliminary)[b], ca. 49+ (finally)[c]	M, partially E

M: manual computation necessary; E: electronic computation possible.
[a] (−) Weak; (+/−) medium; (+) strong; (++) very strong.
[b] Information available after publication of application. Information can still change during the granting procedure.
[c] Lower bound.
[d] Differently computed indicators in different studies.

As of today, however, there exists little theoretical understanding and even less empirical evidence on the complex interaction between indicators, determinants, and prices, costs, and quantities of protected products sold. The complex information hidden in the patent data is still a 'black box' to many economists. It seems particularly puzzling that indicators are not always unambiguously related to a patent's value because they simultaneously reflect complex effects.[2] This paper therefore tries to contribute to a better understanding of interactions and interdependencies between patent value, value determinants, and indicators by analyzing patent attorneys' decisions during the patent application procedure. By doing so, the paper also inspires the compilation of value indicators using so-far unused patent information and enlarges the toolkit of publicly available information to value intangible assets. Before this analysis is undertaken, however, the existing state of the art on the assessment of patents using indicators is briefly summarized in the next section.

2.2. Known indicators of patent value—an overview

Until today, a variety of variables have been tested as indicators of patent value in empirical surveys. Looking at 23 empirical studies related to patent indicators and value, Reitzig (2002a) analyzes the appropriateness of the 13 best-known indicator variables for business purposes. Table 1 summarizes known patent indicators and their advantages and limitations for business purposes. The three columns in Table 1 each refer to one of the evaluation criteria for patent indicators laid out in the introduction to this paper. Column A reports on the validity of the indicator variable. Column B shows the point in time at which the information to compute the indicator becomes accessible. Time is measured in months starting from the

[2] Claims, for example, have been related to the *breadth* of a patent. At the same time, they also reveal
information about its *inventive step* (*non-obviousness*). As *breadth* and *inventive step* may affect the economic value in different ways, however, these ambiguities pose problems on the interpretation of the coefficient of the claims indicator on value. The same problem holds true for several other indicators, especially for those indicators that use highly patent specific-information.

filing date of the patent.[3] Finally, column C reveals whether the information is available electronically or has to be collected manually. All indicators draw from publicly available information. Column A itself is subdivided into two sub-columns that regard the theoretical plausibility and the existing empirical evidence for the validity of the indicator in separate ways. It turns out that forward citations, family size, and the ownership variable show the highest degree of theoretical and empirical validation. However market value also seems to be a good indicator for a company's intellectual property assets.[4] Pioneer work on analysing the relation between backward citations and patent value was carried out by Narin et al. (1997). Forward citations had been introduced by Trajtenberg (1990) and had been validated as indicators of patent value in numerous subsequent surveys, e.g. by Albert et al. (1991), Harhoff et al. (2003), Lanjouw and Schankerman (2001), and Harhoff and Reitzig (2002). Family size—and indicator known from earlier works by Grefermann et al. (1974) and Schmoch et al. (1988)—was introduced as a value indicator by Putnam (1996) and again re-validated by Lanjouw and Schankerman (2001), Harhoff and Reitzig (2002), and Guellec and van Pottelsberghe de la Potterie (2000).[5] The correlation between market value and patents had been examined by Griliches (1981), Conolly et al. (1986), Conolly and Hirschey (1988), Cockburn and Griliches (1988), Megna and Klock (1993), Hall et al. (2000) and Ramb and Reitzig (2004). All the studies mentioned above differ with respect to the quality of the research design, the sample sizes, and the kinds of patents (US, EP, DE). They do, however, have a common feature in that they all validate indicators which are linked to patent value by rationales that speak to rather *general* economic considerations which do not particularly involve in-depth knowledge of institutional details of the patent system. The concept of using citation measures was well known from other disciplines of social science. The fact that ownership affects value is a classical IO consideration. Thus, these indicators may be seen as 'first generation' indicators of patent value. By saying so, no depreciation whatsoever is expressed. On the contrary, the indicators seem reliable and helpful for the evaluation of patents.

In more recent times, other observable information from patent databanks was taken to compile further proxy variables of patent value. In his study, Lerner (1994) successfully linked the market value of 535 biotech companies to the number of patents and the average number of four-digit International Patent Classifications (IPC) of the companies patents. His goal was to operationalize the 'breadth' or the 'scope' of a patent. Unfortunately, the 'scope' variable turns out to be an insignificant regressor in many of the subsequent surveys.[6] Guellec and van Pottelsberghe de la Potterie (2000) and Harhoff and Reitzig (2002) computed further indicators speaking to *patent-specific economic* considerations, such as referring to the filing strategy or the legal contents of backward citations. Obviously, as of today there exists less empirical evidence for these 'second generation' indicators that use patent-specific procedural information and link it to patent value or patent value correlates. Still, the indicators are appealing as they take on the patent-specific knowledge and use it for the computation of value proxies.

Up until now, however, very few researchers have exploited the last resource of information available on patents (i.e. the patent full-text documents themselves). Both 'first and second generation' indicators only make use of 'first page' information stored in databanks. To patent attorneys, this seems strange to some extent, since most of the information on a protected technology and its anticipated economic value is conveyed in the patent draft itself. But then again, special knowledge is required to decipher the relevant information, which is codified in the patent document in a very special kind of way. Tong and Frame (1992) were the first to use information from patent documents and make an attempt to compute what

[3] Note that the information on time is only valid for DE or EP patents.

[4] Note that the 'market value' indicator differs from the other indicators in three respects. First, the marketvalue of a company only allows to serve as an indicator of the aggregate value of intellectual property assets of the company. Besides, many of the empirical studies on the correlation between market value and the number of patents report on a lag structure which has to be taken into account. Finally, 'market value' is information that does not draw from publicly available patent databanks.

[5] See Guellec and van Pottelsberghe de la Potterie (2000).

[6] See, however, Shane (2001) for the importance of patent scope for the analysis of firm creation.

I will call the 'third generation' indicators. They correlated the number of claims in a patent draft to several macroeconomic indices of a nation's technological performance. Most recently, Lanjouw and Schankerman (2001) utilize the information on claims to model the probability of challenge and validity suits for a sample of US patents. The number of claims has been regarded as a possible operationalization of a patent's 'breadth'. Third generation indicators (i.e. any indicators compiled from the patent full-text itself) seem to have one major advantage and one major disadvantage over other indicators. They are attractive since they are available early in time (directly after the publication of the patent) and since they show a strong theoretical foundation. Their disadvantage lies in their endogeneity; i.e. that the patent document is drafted by the proprietor (or his attorney) who therefore has the opportunity to "infer on" the value of his patent by the mode of drafting the document.

Still, when thinking of ways to develop new value indicators, the greatest potential lies with second and third generation variables. The challenge here is to understand the codification of technology and value-related information by patent attorneys in such detail that compilations of new indicators show a maximum of theoretical foundation, and a minimum of ambiguity and endogeneity. The following section therefore sketches the strategic considerations followed by patent attorneys during the filing process and opens up the black box of their codification.

2.3. Expanding the theory—opening the "black box" of patent attorneys' work

Interviews were conducted with nine senior experts from patent law firms, a corporation's patent department, and the European Patent Office (EPO). All of the experts had many years of experience in the field of filing and enforcing European Patents in the field of chemistry or chemical engineering. Despite their current specializations and client profiles, all of the patent attorneys had had many years of experience in working for small, medium, and large clients and entertained close contacts to corporate patent departments. As it turned out, the core of the patent attorneys' work is to maximize profits from legal protection for a given invention. Economically speaking, the patent attorneys' work comes closest to a decision-making problem under uncertainty. I will therefore first outline the decision problem in an abstract way. Then I show how exogenous and endogenous variables (from the standpoint of the patent attorney) enter the attorney's rationale. I will focus mainly on the *state of the art*, the *inventive step*, and the *breadth* of the patent. Since those decision variables are latent variables, I will finally outline how the attorney's rationale translates into observable action. Here, I will focus on the draft of the patent application and briefly mention two procedural steps that have not yet been described in the literature.

The following descriptions of the decision-making problem the attorney faces refer to the European Patent system. Thus, some procedural details cannot be directly transferred to the US system. The basic material trade-offs, however, also hold true for US patents.

Patent protection in Europe can be achieved in three ways. Either the applicant chooses separate national filings in the countries in which he/she seeks protection or he/she decides to file a central European application according to the European Patent Convention (EPC) leading to a European Patent (EP). A third possibility is to use a global priority (PCT) application and subsequently decide for one of the two ways described above. The modes differ with respect to fixed and variable costs. As a rule of thumb it may be stated that the fixed costs of filing increase going from the national, via the European to the global application mode. At the same time, variable costs for additional designated states of protection decrease in the same order. As the data set in this paper is based on patents filed exclusively via the EPC or PCT, the description of the decision-making process is limited to EP patents only. The 'life' of a patent in Europe may take several paths. After its grant it can be centrally legally attacked in a so-called opposition procedure within nine months. Third parties gain the chance to diminish or completely destroy the patent's validity for its entire territory of legal effect. The territory of legal effect is chosen by the patent holder. He/she designates the countries for which he/she seeks protection and incurs variable costs for each country. The EPO decides on the opposition filed and either upholds, amends, or revokes the patent. Appeals against decisions on the opposition plea by the EPO can be filed from either side, the patent holder and the opposing party. Fig. 1 shows the legal 'life-tree' of an EP patent.

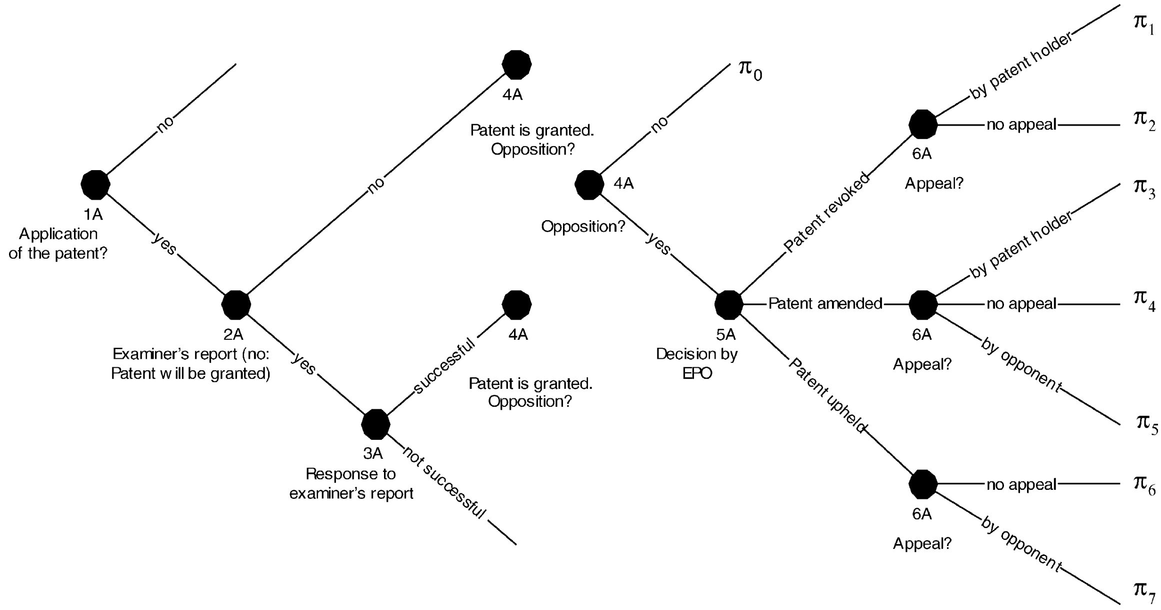

Fig. 1. The life tree of a European Patent (EP).

Using the tree in Fig. 1, the patent attorneys' work can now be described in an abstract way. (Anticipated) Profits can be assigned to all the outcomes of the tree. Probabilities can be assigned to the occurrence of the different legal scenarios (not illustrated in Fig. 1). The patent's value is then the sum over the expected profits (i.e. profits times probability of scenario) in all possible scenarios. It is the job of the attorneys (in cooperation with technology managers) to influence profits in discrete scenarios and probabilities of different scenarios becoming true in such a way that the overall expected profits are maximized.

According to the experts, the *state of the art*, the *inventive step*, and the *market size* underlying the protected invention are the most important exogenous parameters in the maximization process. Besides, the industry often dictates whether the patent can be used as an *exclusion right* in the traditional sense or whether it may rather serve as a *bargaining chip* in technology negotiations with other companies (see Rahn, 1994; Hall and Ham-Ziedonis, 2001). The most important setscrews to be influenced by the patent attorney on the other hand are *breadth*, *disclosure*, and the *mode of filing*. What makes the maximization process complex is that the endogenous variables influence the patent's overall expected value in opposite ways through the probabilities and the static profits. In fact, trading off between the different effects of the endogenous variables is therefore a crucial part of the attorneys' work as will become clear from the following.

At the first meeting between patent attorney and inventor, the expected net profits from protecting the invention are assessed, basing the estimation on the exogenous parameters mentioned above. The estimations are very qualitative, but this is how the attorneys value the exogenous variables:

- Little state of the art hints at maximum at a 'latent' market where benefits from patenting can be expected in the future.
- Comprehensive state of the art points at an active market and patenting seems profitable. However, an increasing state of the art raises the risk of legal conflict with competitors and therefore decreases the expected profits.
- If *inventive step* is small and there is little *state of the art*, expected profits are small.
- If *inventive step* is small and there is comprehensive *state of the art*, possible profits are high. However, the risk of losing the patent in a legal argument rises, too, decreasing the overall expected profits

from patenting. Expected profits may range from medium to high.

- If *inventive step* is high and there is comprehensive *state of the art*, possible profits seem high, and there is little risk of losing the patent in a legal argument. Expected profits are very high.

Given the exogenous variables, the patent attorney can maximize profits by adjusting the endogenous variables with respect to the situation he/she is facing. He/she will extend the *breadth* to its maximum for patents showing a high *inventive step* and possibly high profits. By doing so he/she maximizes the profits for each scenario in Fig. 1. He/she may well increase the probability of a legal attack at the same time, but the probability of losing in the opposition case is small. The fixed costs for the opposition are outweighed by the increase in the profits. Conversely, the attorney will reduce the *breadth* for patents with a decreasing *inventive step*. The higher the possible profits from a valid patent the more he/she will reduce the *breadth* given the same *inventive step* since he/she does not want to lose the patent in a legal dispute. The considerations are similar though slightly different for bargaining chip patents. Here, legal disputes are the exception and the attorneys will only make sure that the application 'survives' the granting procedure.

Until this point, consideration was only given to the latent variables that drive the rationale of patent attorneys. The attorneys' considerations, however, manifest themselves in the patent draft. Thus, by looking at the patent draft, it should be possible to gain hints at (and ultimately indicators of) the anticipated value of the patent by the attorneys. The interviews reveal that this task may in practice be aggravated by the fact that different patent attorneys have individual *modes of drafting* and that considerable noise should be expected when pursuing a patent text analysis. Still, in principle the following passages in the patent draft should reveal the information of interest:

- The *state of the art* is described in the first section of the patent.
- The degree of *inventive step* is reflected in the description of the technical problem. The technical problem is normally presented following the description of the *state of the art*. Its solution is presented in the *disclosure* of the patent, and summarized in the
- Claims' section at the end of the patent. Claims also refer to the *inventive step* behind the patent. At the same time, the *breadth* of the patent should be reflected in the claims. In the chemical industry, especially the number of independent product claims should be an indicator of patent *breadth*.
- Dependent product claims, process- and application claims also add to the *breadth* of the patent. At the same time they operationalize what patent attorneys call fall-back options for legal disputes. Their number should rise with an increasing risk of legal attack (falling *inventive step*, increasing profits in scenarios).
- Finally, technical advantages and preferred technical solutions in the *disclosure* should also serve as hidden fall-back options. On the other hand, they often demonstrate that inventor and attorney already have an application of the invention on their mind, pointing at an existing market.

As mentioned above, the decision-making process of the attorneys takes place under uncertainty. Thus, before drafting the patent application, attorneys will try to gather as much information about the underlying *state of the art* and the market size as they can. The information will *ceteris paribus* enhance their ability to assess the patent's *novelty* and *inventive step* and hence its economic value. A way to gather information more quickly than usual is to request an accelerated search report on the state of the art from the EPO. A way to "buy" decision time is to file the patent through the PCT.[7] Once attorneys decide that protection is valuable and should be acquired as soon as possible they can accelerate the granting procedure in the European system by requesting an accelerated examination. On the global level, they can accelerate protection by applying through the so-called chapter II of the PCT.[8]

3. The empirical research design

To validate new indicators of patent value, this paper attempts to link patent value to observable procedural information and to the design of certain text passages in the patent draft in a large-scale empirical study. As

[7] Further details follow in the interpretation of the multivariate statistical results.

[8] See footnote no. 7.

valuations of patents are very hard to get, a patent value correlate is chosen as the dependent variable in the regressions, namely the likelihood of an opposition against the patent. In the following, I will briefly sketch why the approach seems plausible in general but I will also point at the interpretation problems of the regression results that occur from the chosen design.

3.1. Opposition and patent value

Extending the model by Lanjouw and Lerner (1997) and Harhoff and Reitzig (2002) can show that the condition for the occurrence of an opposition is given by formula (1):

$$P = 0 \quad \text{if } j\alpha \geq \frac{jW - (L + l) + S}{w} \qquad (1)$$
$$P = 1 \quad \text{o.w.}$$

In formula (1), $j\alpha$ corresponds to the value of the valid patent for its owner and j is the benefits of a successful opposition for the opponent. W is the anticipated probability by the opponent of winning the opposition, w is the patent holder's anticipated probability of losing the patent. L and l refer to the litigation costs for both parties, and S are the settlement costs. Formula (1) illustrates that the probability of an opposition is correlated with the value of the valid patent for the patent owner, $j\alpha$. This observation supports the research design chosen in this study. At the same time, however, formula (1) also shows that the likelihood of an opposition depends on probabilities of the opposition outcome as anticipated by the opposing parties. In fact, if settlement costs exceed litigation costs by large, the settlement option becomes negligible and the likelihood of an opposition is described by formula (2):

$$P = 0 \quad \text{if } j \geq \frac{L}{W} \qquad (2)$$
$$P = 1 \quad \text{o.w.}$$

Thus, when interpreting regression results of the likelihood of an opposition on indicators, two things should be kept in mind. At first, the likelihood of an opposition is driven by the profits j of the *opponent* in the case of a successful opposition. Those should be highly correlated though not necessarily identical to the value of the valid patent for the *owner*, $j\alpha$. In a simple one-product world where the patent protects a single product and there are only two players, j would be the duopoly profits of the *opponent* whereas $j\alpha$ would be the monopoly profits of the patent *owner*. To facilitate the following descriptions, I will refer to the opponent's benefits from a successful opposition as the *patent's value*. Assuming that j and $j\alpha$ are similar, the patent's absolute value will be similar for both, the patent owner and the opponent.[9] Secondly, proxy variables may well refer to both the value of a successful opposition for the opponent as such, and the anticipated probability of the outcome of the opposition procedure.[10]

3.2. Data collection and computation of indicators

The only available source of EP patent full-texts in machine-readable format at the time of the study was the EUROPATFULL© databank maintained by the Fachinformationszentrum Karlsruhe/Germany. At the day of the data collection, the full-text patent data were only available for EP patents granted between 1992 and 2000. Given the average time of around 4.3 years for granting a patent at the EPO, I chose patent filings for the years 1992–994 for the study. I decided

[9] For a discussion of the opponent's incentives to file, an opposition to enhance his bargaining position withthe patentee, see Harhoff and Reitzig (2002) for a comprehensive discussion.

[10] As pointed out by one of the referees of this paper, an alternative estimation using the number of opponentsas the dependent variable appears intuitive in that it might have the advantage of introducing more variation on the left hand side of the regression. From a theoretical standpoint, an increased number of opponents may hint at a more competitive scenario for the patentee. While, on the one hand, this may indicate an increased value of the *upheld* patent for its owner (because the discrepancy between the patentee's profits and his competitors' profits—ergo the asset value of the patent—rise if the patentee wins the opposition), on the other hand the likelihood of the patent being withdrawn may increase, too (because more competitor resources are put into its revocation). Thus, the use of a dependent variable with more variation (number of opponents) again comes with a cost. Given this fact, and given that there has not yet been developed a formal model that establishes a clear link between the number of opponents and the value of a patent, this papers uses the binary occurrence of an opposition as the dependent variable despite its lower variation. The reader may, however, be interested in the fact that the multivariate results presented in Table 5 do not change at all with respect to the significance of individually significant coefficients when running an alternative ordered probit regression that differs from regressions 5A, 5B, and 5C only in that it uses the number of opponents as the dependent variable! This finding (not reported in the paper in detail) insinuates that the number of opponents should reflect the degree of competition in future modelistic formalizations.

to focus on patent filings from the chemical industry so that the patent rights would be exclusion rights rather than bargaining chips (see, e.g. Cohen et al., 2000). The sampling was based on a four digit IPC classification for industries as proposed by Schmoch and Kirsch (1994). The patents come from six different chemical branches: organic fine chemicals (37%), polymer chemistry (38%), pharmaceuticals (5%), biotechnology (3%), agricultural chemistry (1%), and petroleum chemistry (16%). As the computation of indicators referring to the wording of the patent draft should originally be carried out using a text scanning software for German language, I also decided to look only at patents who would have a German, Swiss, or Austrian inventor.[11] Out of a sample of 2570 remaining patents 1000 documents were chosen randomly, the only alteration being to ensure that the opposition rate would be on the same order of magnitude as the long-term average rate of 8.1% for all EP patent.[12] Outlier correction finally brought the sample down to 813 observations.

Using four further databanks, indicators drawing from procedural patent information were computed. Among the indicators listed in Table 2 are 10 indicators (the first 10 in the table) that have been used in earlier studies. The last two indicators refer to the patentee's options to accelerate the production of the search report by the EPO or to accelerate the granting procedure at the EPO that have not been tested before. Table 2 lists the variables in the first column and reports on their computation algorithm in the second column.

Contrasting Tables 1 and 2, it is conspicuous that forward citations are not computed in this study. The rationale behind this is very straightforward. The obvious disadvantage of forward citations as value indicators is their late availability in time. The goal of the paper, however, is to present indicators that serve as proxy variables for patent value at an early stage of the patents' life.

Aside from the indicators utilizing procedural patent information, I computed variables that directly draw from the full-text of the patent draft.[13] The indicators are described in Table 3. The first column of Table 3 names the indicator and the second one briefly recalls the link between the indicator and the economic value of a patent.

3.3. Derivation of hypotheses

Column three of Table 2 and column five of Table 3 show the expected signs of the variable coefficients when tested as correlates of the likelihood of an opposition. Whereas correlations between procedural indicators and the likelihood of an opposition should be primarily mediated via the patent's value, there seems to be a more complex relation between the text indicators and the likelihood of the opposition. Columns three and four in Table 3 therefore distinguish between the expected sign of the correlation between the profits from the protected invention and the text indicators, and the anticipated probability of the opponent to win in an opposition and the text indicators respectively. Column five of Table 3 then provides a very preliminary expectation of the aggregate effect that text indicators should have on the likelihood of an opposition.

Due to possibly counteracting effects associated with the text indicators and resulting ambiguities concerning specific expectations in the empirical study, I chose to test two hypotheses of reduced information contents in this paper. The hypotheses tested in this paper are the following:

[11] The software searches for German keywords in the patent draft and can therefore only be applied to German documents for the time being. First results showed, however, that for the time being the software still yields significantly different results from a manual compilation which is why the indicators were computed manually for this study once more.

[12] Note that the specific data constraints to construct the sample for the analysis in this paper (language constraints due to the full text analysis) led to an inevitable shit in the distribution of patents over industries compared to the total number of EP patents filed between 1992 until 1994. When looking at the total number of EP patents filed in the respective IPCs during that period, the following distribution is found: organic fine chemicals (32%), polymer chemistry (25%), pharmaceuticals (15%), biotechnology (9%), agricultural chemistry (5%), and petroleum chemistry (14%). Since opposition rates across these industries differ, the discussion of the multivariate estimation results (Tables 5 and 6) based on the current sample must account for the potential selection bias due to the sample's particular industry distribution. Consequently, the specifications in Tables 5 and 6 contain industry dummies that filter out any specific industry effects that would most likely have an impact on the estimation results and in turn render the discussion more complex.

[13] Note that the potential operationalizations were discussed in depth with some of the patent experts. The author wishes to thank Dr. Thomas Koch, European Patent Attorney (Hoffman & Eitle) in particular for discussing respective ideas.

Table 2
Indicators computed from procedural information

Variable	Algorithm	Expected effect on opposition
Backward citations to patent literature	Number of patent references to the state of the art that are actively quoted by the patent	+
Backward citations to non-patent literature	Number of non-patent references to the state of the art that are actively quoted by the patent	No prediction (various effects possible)
Family size	Logarithm of the number of designated states	+
'Scope'	Number of four-digit IPC classes	+
PCTI application	Dummy variable taking on the value 1 if the patent was filed via PCT and the period of time between filing date and entry into the regional phase is 20 months or less	+
PCTII application	Dummy variable taking on the value 1 if the patent was filed via PCT and the period of time between filing date and entry into the regional phase exceeds 20 months	+
Share of A-classifications among backward citations	Share of backward citations among the total number of backward citations which were considered relevant but not innocuous by the preliminary examiners in the Hague	–
Share of X-classifications among backward citations	Share of backward citations among the total number of backward citations which were considered potentially innocuous by the preliminary examiners in the Hague	+
Number of inventors	Total number of inventors	(+)
Number of applicants	Dummy variable taking on the value 1 if more than one applicant is mentioned in the patent	(−)
Accelerated search request	Dummy variable taking on the value 1 if a request was filed for an accelerated production of the search report	(+)
Accelerated examination request	Dummy variable taking on the value 1 if a request was filed for an accelerated examination	+

H1. There is a correlation between yet untested procedural patent indicators (such as acceleration requests) and the likelihood of an opposition.

H2. There is a correlation between yet untested text indicators and the likelihood of an opposition.

Note that the two hypotheses cannot reflect all the potential trade-offs elucidated in the theoretical section above. The study presents a first test of partly new indicators and the hypotheses are consequently kept as simple as possible.

4. Empirical results and discussion

In the following, the empirical results are presented and discussed.

4.1. Descriptive statistics

Table 4 summarizes the data as it was used in the study. The upper part of the table refers to procedural explanatory variables, the lower part shows the means for the text indicators.

As can be seen from Table 4, about 13% of the patents in the sample were opposed. Thus, the opposition rate in the sample is slightly higher than the long-term average of 8.1% for the industry. Patents in the sample might therefore be a little more valuable than on average for this industry. Moving on to the other procedural information, some peculiarities can be observed. Whereas the number of inventors and applicants seem very plausible comparing them to earlier studies, the PCT application ratios appear to be quite low. In fact, further cross checks of the data with the official bulletin issued by the EPO (1998) confirm

Table 3
Indicators computed from full-text information

Variable	Link/operationalization	Expected effect on patent value	Expected effect on opponent's anticipated probability of winning	Expected observable effect on likelihood of opposition
Number of words describing the state of the art	*State of the art/novelty*	+	Unknown[a]	(+)[b]
	(*Disclosure*)	+	−	
Number of words describing the technical problem	*Degree of inventive step*: Problem of the protected invention as 'counterpart' to the technical solution and therefore to the degree of inventive step	+	−	(+)
Number of mentioned technical advantages of the invention	*Expected demand:* Technical advantages as a sign of product- and market proximity	+	Unknown	(+)
	Technical advantages as hidden fall-back options	Unknown	+	
Number of technical preferences of the invention	*Expected demand:* Technical preferences as a sign of product and market proximity	+	Unknown	(+)
	Technical preferences as hidden fall-back options	Unknown	+	
Number of independent claims	*Degree of inventive step*: Independent claims as concise description of the solution and therefore of the degree of inventive step	+	−	(+)
	Breadth	+	Unknown	
	Expected demand: Product and market proximity	+	Unknown	
Number of dependent claims	See above (independent claims)	+	(−)	(+)
	Fall-back options	Unknown	+	
Number of process claims[c]	See above (independent claims)	+	(−)	(+)
	Fall-back options	Unknown	+	
Number of application claims[c]	See above (independent claims)	+	(−)	(+)
	Fall-back options	Unknown	+	

[a] In these cases, the existing theory does not allow to predict an effect.
[b] Effects in brackets are subject to greater uncertainty.
[c] Process and application claims may also be independent claims. To ensure a better distinction between independent product claims and process- and application claims, however, I decided to count them separately in this study.

that the low percentage can be attributed to the selection criteria of the sample. With respect to the requests for accelerated search or examination, the means in the sample again correspond to the long-term average value for EP patents across industries and seem therefore plausible. Going further down in Table 4, some observations seem noteworthy when looking at the text indicator variables. At first, all of the explanatory variables show remarkable variation which is intuitively positive. With respect to the number of independent claims, the mean value of 0.64 deserves some explanation. In order to distinguish between independent product claims and other independent claims (process or application claims), I counted product claims separately. Thus, the number of independent claims only refers to product claims, so does the

Table 4
Descriptive statistics

Variable	Mean	S.D.	Minimum	Maximum
Opposition (0: no; 1: yes)	0.13		0	1
Family size (ln (1 + number of designated states))	2.12	0.45	0	2.83
Number of inventors	3.35	1.66	1	8
Number of applicants >1 (0: 1 application; 1: >1 application)	0.02		0	1
"Scope"	2.24	1.87	1	13
PCTII application (0: no; 1: yes)	0.03		0	1
PCTI application (0: no; 1: yes)	0.01		0	1
Accelerated examination request (0: no; 1: yes)	0.02		0	1
Accelerated search request (0: no; 1: yes)	0.01		0	1
Backward citations to patent literature[a]	3.52	2.03	0	13
Backward citations to non-patent literature[a]	0.81	1.30	0	12
Share of citations classified as "A" among the total number of backward citations[a]	0.46	0.36	0	1
Share of citations classified as "X" among the total number of backward citations[a]	0.12	0.25	0	1
Number of words describing the state of the art	326.07	257.19	0[b]	2115
Number of words describing the technical problem	36.20	29.99	0[c]	295
Number of technical advantages	4.41	5.69	0	49
Number of technical preferences	35.74	37.72	0	304
Number of independent claims	0.64[d]	0.58	0	7
Number of dependent claims	2.51	3.25	0	21
Number of process claims	4.28	4.15	0	26
Number of application claims	0.77	1.50	0	11

N=813.

[a] The marginal difference in the mean of the variable compared to Reitzig (2002b) is explained by the fact that this paper (other than Reitzig, 2002b) also counts references as EP backward citations to the (non-) patent literature that were made by the EPO in its function as one of WIPO's International Search Authorities. The difference in the two different declarations of the variable has only marginal consequences for the multivariate statistics. The significances of individual coefficients do not change from Reitzig (2002b) to this paper; effect sizes for the PCTII indicator and the patent citations decrease respectively increase slightly. The author thanks Dietmar Harhoff for pointing him at this phenomenon.

[b] In these cases, it was impossible to unambiguously identify a passage in the text that was solely referring to the state of the art (see text).

[c] In these cases, it was impossible to unambiguously identify a passage in the text that was solely referring to the technical problem (see text).

[d] As independent process- and application claims were counted separately from independent product claims, the mean of 'independent' (product) claims may well be below unity.

number of dependent claims. Process and application claims were counted separately, but here no distinction between dependent and independent claims was made.

4.2. Multivariate analysis

Seven different models of the likelihood of an opposition on value indicators are shown in this section. Table 5 presents three estimations (downward testing) of the likelihood of an opposition using a simple probit model based on indicators that draw from both known and untested procedural patent information.[14] Besides, dummy variables for the separate chemical branches enter to ensure that industry effects are not attributed to explanatory power of the indicators.

[14] Note again that despite their obvious importance forward citations are not inserted in the specification for a good reason: they are not available until quite some time after patent grant. Since this paper aims at identifying value correlates that can be applied to both old and young patents, forward citations are excluded from the analysis.

Table 5
Likelihood of opposition modeled by procedural indicators (simple probit) (A: full specification; B: downward tested specification; C: marginal effects of B)

Independent variable	Column A: probit coefficient (S.D.)	Column B: probit coefficient (S.D.)	Column C: coefficient of marginal effect in regression B (S.D.)
Family size	0.34** (0.14)	0.35*** (0.13)	0.06*** (0.02)
Number of inventors (coefficient × 10)	−0.56 (0.42)	–	–
Applicants >1	0.13 (0.43)	–	–
'Scope' (coefficient × 10)	−0.07 (0.37)	–	–
PCTII application (0: no; 1: yes)	2.66*** (0.45)	2.67*** (0.50)	0.82*** (0.08)
PCTI application (0: no; 1: yes)	−0.05 (0.52)	–	–
Accelerated examination request (0: no; 1: yes)	1.30*** (0.498)	1.28*** (0.41)	0.39*** (0.16)
Accelerated search request (0: no; 1: yes)	−0.30 (0.55)	–	–
Backward citations to patent literature (coefficient × 10)	0.90*** (0.29)	0.90*** (0.28)	0.16*** (0.05)
Backward citations to non-patent literature	−0.12* (0.07)	–	–
Share of citations classified as "A" among the total number of backward citations	−0.00 (0.19)	–	–
Share of citations classified as "X" among the total number of backward citations	0.21 (0.27)	–	–
Dummy for organic fine chemistry	−0.89** (0.46)	−0.54*** (0.15)	−0.09*** (0.02)
Dummy for polymer chemistry	−0.54 (0.45)	–	–
Dummy for pharmaceutical chemistry[a]	−0.30 (0.51)	–	–
Dummy for biotechnology[a]	−0.32 (0.55)	–	–
Dummy for petrol industry			
Basic chemicals	−0.31 (0.45)	–	–
Constant	−1.53*** (0.54)	−2.15*** (0.29)	–
Wald χ^2 (A: 17/B: 5/C: 5)	105.31 ($P < 0.001$)	72.53 ($P < 0.001$)	72.53 ($P < 0.001$)
Pseudo R^2	0.23	0.21	0.21
N	813	813	813

[a] *Note*: In Reitzig (2002b), the industry dummies for pharmaceutical chemistry and biotechnology were confused with one another. The correct distinction is presented above. Since both variables serve as control variables only and as their coefficients are almost always insignificant in the regressions, however the mistaken declaration in Reitzig (2002b) is of minor relevance.

* Significant at 10% level (two-tailed tests).

** Significant at 5% level (two-tailed tests).

*** Significant at 1% level (two-tailed tests).

Specification 5A is significant at the 0.01% level. Besides, six variables are individually significant, namely the family size,[15] the number of inventors, the PCTII indicator, the indicator for an accelerated examination request, backward citations to the patent and non-patent literature, and the dummy variable for organic fine chemistry. Looking at the PCT indicators, it seems interesting that a filing according to chapter II of the PCT is highly correlated with the likelihood of an opposition whereas the PCTI indicator is insignificant. The results cannot be explained by existing rationales that simply discuss the applicants' *ex-ante* willingness to incur increased filing costs for global protection. In deed, the argument explaining the empirical results is more complicated. Applicants filing patents through PCTII may choose the option for two different reasons that are properly opposed to each other. Either they are initially very uncertain about the economic success of the patent's underlying invention and they choose the PCTII option to "buy" additional decision time over PCTI (by delaying cost intense decision during the filing procedure for a relatively small "option premium") or, on the contrary, the economic success of the patent's underlying invention is free of doubt already at the date of filing and PCTII is used to seek global protection as fast as possible. The fact that PCTII filings may be used for

[15] Note that the coefficient does not change significance if the variable is constructed as a threshold variable or not.

these opposed reasons explains the apparent regression paradoxon in this paper. Only *ex-post*, namely in a sample of granted patents like the one used for this analysis, is the PCTII indicator a true correlate of patent value. In such a sample of granted patents, however, PCTI is only an indicator of uncertainty at the beginning of the filing process and no significant correlation with patent value can be expected.[16] Ending the discussion on the PCT indicators, one observation seems worthwhile to recall. The current sample comprises patents with (application) years between 1992 until 1994. During the recent past, however, the total number of PCT filings at the EPO has increased remarkably. Thus, it appears reasonable to assume that the PCTII indicator might lose some of its explanatory power when testing it on samples of more recent patents.

Finally, the differences in the levels of significance for the two acceleration requests need to be explained. While the accelerated search request does not involve any costly decision or commitment by the applicant, the accelerated examination request involves a cost commitment by the patentee. It is therefore not surprising that the last indicator is significantly correlated with the likelihood of an opposition whereas the first one turns out to be insignificant. Column B of Table 5 shows a specification in which the individually and/or jointly insignificant coefficients of specification 5A were dropped.[17] By showing the marginal effects of specification 5B, column 5C conveys an impression of the orders of magnitude of the different effects. The strongest effects are for the PCTII indicator. PCTII applications in this sample are 82% more likely to be involved in an opposition than other patents. Also the accelerated examination request indicator is very strong. Patents which were examined in an accelerated procedure are roughly 40% more likely to be attacked by opposing parties than other patents. It may be concluded that hypotheses H1 is preliminarily confirmed by the data.

Having discussed the procedural indicators, the next three regression models are based on text indicators. The estimations use a probit model with correction for heteroscedasticity as proposed by Harvey (1976). I chose the heteroscedastic probit model for two reasons, a theoretical and a statistical one. Theoretically, the interviews with the patent attorneys pointed at the problems of differing *modes of drafting* patent applications leading to systematic noise in the data across various applicants.[18] Statistically, models 6A through 6C support this assumption.[19]

Table 6 shows the results of three different regressions (downward testing) of the opposition variable on text indicators.

Column 6A models the likelihood of an opposition using all text indicators computed in the study. The upper part of Table 6, column A, shows the first regression results of the most comprehensive specification in which all explanatory variables are used to model the likelihood of an opposition and the *variance* of the dependent variable at the same time. The lower part of Table 6, column A, shows the respective auxiliary regression. Here, the coefficients describe the correlation between the explanatory variables and the variance in the main regression. Following an approach published by Lechner (1991), I carried out joint tests of significance for the individually insignificant coefficients in the auxiliary regression 6A to arrive at a robust specification for the auxiliary regression containing only

[16] The results confirm in part the (pioneer) findings by Guellec and van Pottelsberghe de la Potterie (2000), who also find that PCTII should indicate a higher value of a patent application than PCTI. However, the authors give no complete explanation for the phenomenon they observe.

[17] Note that the coefficient for the number of citations to the non-patent literature is significant at the margin in specification 5A. A joint χ^2-test with the individually insignificant variables in model 5A ($\chi^2(12)$-test: 9.66; $P = 0.65$), however, confirms that the variable is not an important predictor of the likelihood of an opposition in the final model. The counterintuitive finding in 5A (negative coefficient) may in part be due to the noise of this variable. Currently, databases still do not filter out "patent-related" citations from the non-patent literature section (e.g. Japanese Patent Abstracts or Derwent Patent Abstracts). This observation is consistent with Meyer (2000) who notes that there exist various types of non-patent citations, and not all of them reflect references to basic research.

[18] Note that there may be various reasons for the different modes of drafting patent applications. One variation may originate from the individual character of actually formulating technical problems, solutions, and claims by different attorneys within the confines of their legal context. Another source of variation may come from the different requests by the applicant himself (e.g. certain applicants will put more emphasis on disclosing as little technical knowledge as possible than others, etc.).

[19] Not surprisingly, simple probit specifications (not shown in this paper) do not yield statistical evidence for the suitability of the full-text indicators.

Table 6
Likelihood of opposition modeled by text indicators (heteroscedastic probit) (A: full specification; B and C: downward tested specifications)

Independent variable	Column A: coefficient in the main regression (S.D.)	Column B: coefficient in the main regression (S.D.)	Column C: coefficient in the main regression (S.D.)
Number of words describing the state of the art (coefficient × 1000)	−0.21 (0.16)	−0.22* (0.12)	−0.26 (0.20)
Number of words describing the technical problem (coefficient × 100)	−0.08 (0.22)	0.29 (0.19)	0.38* (0.21)
Number of technical advantages (coefficient × 10)	0.08 (0.11)	0.10 (0.07)	–
Number of technical preferences (coefficient × 100)	−0.43 (0.37)	−0.88* (0.54)	−1.63** (0.72)
Number of independent claims	0.16*** (0.05)	0.13*** (0.04)	0.14*** (0.04)
Number of dependent claims (coefficient × 10)	0.61*** (0.17)	0.53*** (0.10)	0.64*** (0.13)
Number of process claims (coefficient × 100)	−4.10* (2.18)	−2.86** (1.34)	–
Number of application claims (coefficient × 10)	0.83*** (0.27)	−0.11 (0.32)	–
Dummy for organic fine chemistry	−1.22*** (0.34)	−0.96*** (0.27)	−0.81*** (0.18)
Dummy for polymer chemistry	−0.42* (0.25)	−0.27 (0.22)	–
Dummy for pharmaceutical chemistry	−0.03 (0.27)	0.09 (0.23)	0.31* (0.17)
Dummy for biotechnology	−0.39 (0.33)	−0.31 (0.29)	
Dummy for petrol industry	−0.21 (0.24)	−0.06 (0.21)	–
Constant	−0.57** (0.27)	−0.61*** (0.22)	−0.91*** (0.14)
Wald χ^2 (A: 13/B: 13/C: 7)	43.52 ($P < 0.001$)	104.30 ($P < 0.001$)	73.94 ($P < 0.001$)
N	813	813	813
Auxiliary regression ($\ln \sigma^2$ dependent variable)	Coefficient	Coefficient	Coefficient
Number of words describing the state of the art (coefficient × 1000)	0.08 (0.26)	–	–
Number of words describing the technical problem (coefficient x 100)	0.28 (0.31)	–	–
Number of technical advantages (coefficient × 10)	−0.02 (0.15)	–	–
Number of technical preferences (coefficient × 100)	0.47* (0.29)	0.81*** (0.30)	1.03*** (0.31)
Number of independent claims	−0.27** (0.14)	−0.31*** (0.11)	−0.32*** (0.10)
Number of dependent claims	−0.10*** (0.03)	−0.10*** (0.04)	−0.05** (0.02)
Number of process claims (coefficient × 10)	0.28 (0.24)	–	–
Number of application claims (coefficient × 10)	−1.64** (0.84)	–	–
Likelihood ratio test for $\ln(\sigma^2)$: χ^2 (A: 8/B: 3/C: 3)	25.74 ($P < 0.001$)	21.58 ($P < 0.001$)	21.21 ($P < 0.001$)

* Significant at 10% level (two-tailed tests).
** Significant at 5% level (two-tailed tests).
*** Significant at 1% level (two-tailed tests).

few variables.[20] The result is presented in the lower part of column 6B. The upper part of column 6C finally presents a specification modeling the likelihood of an opposition with correction for heteroscedasticity using only few variables.

[20] Again, the approach reflects the experimental character of the study. Even though the theoretical discussion in the preceding section would also justify more complex ways of defining specifications and constructing variables, e.g. interacted terms, I stick with the simplest testing procedure in this first validation study.

As can be seen from column 6C, four of the text indicators correlate significantly with the likelihood of an opposition when correcting for heteroscedasticity, namely the number of words describing the technical problem, the number of technical preferences, the number of independent product claims, and the number of dependent product claims. For a variety of reasons mentioned above, these results should be interpreted carefully. Still, the findings are very plausible given the state of knowledge developed in this paper. First of all, it was carefully predicted in Table 3 that the likelihood of an opposition should increase with the length of the problem description. Assuming that the indicator operationalizes the degree of *inventive step*, the length of the problem description will correlate positively with the patent's value. This has a positive effect on the likelihood of the opposition. On the other hand, the opponent's expectations of winning the opposition case should fall with rising *inventive step*. Thus, the result in 6C suggests the following: the positive effect on the likelihood of an opposition due to increased patent value exceeds the negative effect on the opposition likelihood due to the opponent's diminished expectations of winning the case. Hence, the number of words describing the technical problem mainly correlates with the potential profits from protecting the invention.

Moving two coefficients down in column 6C, there is a positive correlation between the number of independent claims and the likelihood of an opposition. Again, the result seems plausible. Assuming that independent product claims operationalize the *breadth* of the patent as suggested in Table 3, then profits from the patent's value should rise with the number of independent claims. If, as also suggested in Table 3, independent claims are also a measure of the *inventive step*, then the opponent's expectation of winning the opposition case should fall at the same time. Hence, the result found in 6C suggests that the same rationale applies to independent claims as to the number of words describing the technical problem. The positive effect on the likelihood of an opposition due to increased patent value exceeds the negative effect on the opposition likelihood due to the opponent's diminished expectations of winning the case. Hence, the number of independent claims should correlate with the patent's value.

Theoretically, the last rationale could simply be applied to the *dependent* product claims, too. However, a more elaborated line of argument seems more convincing. As proposed in Table 3, dependent claims often serve as so-called 'fall-back' options. It seems plausible to assume that patentees are more likely to insert those costly fall-back options into a patent application when they face a higher chance of being attacked. Again, they do face a higher chance of being attacked the more valuable the patent and the higher the opponent's anticipated probability of being successful in an opposition. Given the fact that the patent attorney faces additional costs for inserting each additional dependent claim, there is good reason to believe that also dependent claims are a valid correlate of a patent's value.

Finally, I would have expected the likelihood of an opposition to rise with the number of technical preferences. Table 3 suggests that technical preferences either serve as fall-back options or that that they reflect market proximity. However, in both cases the likelihood of an opposition should not drop. Hence, the results found in 6C conflict with the theoretical expectations and additional explanations are needed. In fact, when looking carefully again at the raw data in detail I find the highest number of technical preferences in those cases where the description of the technical invention is most comprehensive. This leads to the assumption that technical preferences possibly operationalize the *disclosure* of the patent rather than anything else. Assuming that technical preferences are an indicator for the *disclosure* of the patent, the results become then very plausible with respect to Article 83 EPC. According to this article, opponents may substantiate their opposition by blaming the patentee of insufficient *disclosure*. Then, the result in 6C suggests that the patentee invested additional time in the draft of the patent and in carrying out further experiments to reduce the likelihood of a substantiated opposition. His/her willingness to incur extra costs point at a high expected value of the patent. Apparently, however, the negative effect on the likelihood of an opposition mediated by decreased expectations of the opponent of winning in the opposition outweighs the positive value effect on the likelihood of the opposition. Overall, it can be concluded that the data provide some preliminary empirical evidence for the validity of hypothesis H2.

5. Conclusions and outlook on future research

Recalling the scope of the paper, I finally ask as to what extent the paper enhances our understanding of measuring patent value with indicators for corporate purposes. This paper does not claim to solve the applied problem of valuing patent portfolios from a corporate perspective. Neither does it offer answers to various open questions, e.g. on synergetic effects of patents within portfolios, nor would the empirical study contain a test that validates the results across industries. Irrespective of these shortcomings, however, this paper claims to contribute to a better understanding in three different ways. First, it expands the theory by analyzing patent attorneys' work which appears crucial to interpret observable legal actions as indicators of patent value. In particular, new ideas for the compilation of various new indicators arise from the descriptions of patenting rationales. The paper suggests that operationalizations of key variables in the filing process, such as the *state of the art of existing technology*, the *inventive step*, and the *breadth* of the patent should be suitable value indicators. Secondly, the paper confirms convincingly the validity of a new procedural indicator of patent value, the so-called accelerated examination request. The data provide empirical evidence that patents appear more valuable when patentees are willing to make a *cost commitment* early during the filing procedure. At the same time it reconfirms the validity of other procedural indicators of patent value (PCTI and PCTII), however, offering more detailed explanations for their validity than described in the literature so far. Third, the paper provides some preliminary results on the appropriateness of text indicators as additional measures of patent value. As typical for a first and experimental study, the empirical part of the paper could not test the complexity of all the potential dependencies between the full text patent indicators and the likelihood of an opposition. Future studies may take account of the different trade-offs by testing the indicators structurally, or by introducing additional interacted variables. This study is a first and preliminary test of full-text indicators and the results consequently have experimental character. They suggest, however, that new indicators of patent value can be found by a full-text analysis of the technical problem, the number of technical preferences, independent and dependent product claims, and application claims. Further results of this exploratory study (not reported in this paper) confirm that procedural indicators and text indicators are partly complementary in their explanatory power[21], i.e., however, that the new full-text indicators may potentially serve as the desired substitutes when 'classic' indicators, for example citations and family size, do not show variations across the portfolio of a *single* firm. On the other hand, the computation of the full-text indicators appears useful even if established procedural indicators are available, as the two are no perfect substitutes. Finally, two further remarks are of high interest: first, all of the tested indicators in this study are available early in the life-time of a patent, draw from publicly available information and are computable at low cost. Secondly, a joint specification consisting of procedural and text indicators[22] predicts the occurrence of an opposition correctly more than 90% of all cases. Table 7 compares predicted and real opposition occurrences for the whole sample. With respect to the initial question, namely to what extent the paper helps to improve patent valuations from an applied perspective, an interesting quantitative figure can be derived: in this paper the estimated value of a patent portfolio using indicators deviated from the 'real' value of the portfolio by only 5%.[23]

Table 7
Predicted vs. real opposition cases ($N = 813$)

	No opposition (predicted)	Opposition (predicted)
No opposition (real)	706	4
Opposition (real)	76	27

[21] Tested together in one specification, application claims add to the explanatory power of procedural indicators. The results are not presented, though, since the necessary correction for heteroscedasticity among the procedural indicators renders the joint test difficult.

[22] See footnote 21. The specification consisted of the family size, the PCTII indicator, the accelerated examination request, and the number of application claims. After downward testing, the specification was found to be the most efficient parameterization and hence used for the final applied prediction example.

[23] The figure was derived as follows: repeated random samples (10 times, consisting of roughly the half of the total portfolio of 813 patents) were chosen and predictions of opposition occurrences were calculated. Starting from the premise that opposed patents are on average more valuable than opposed patents by two orders of magnitude, the *predicted* values for portfolios of the randomly chosen sub samples were calculated by summing up the values of the individual patents and then compared to the '*real*' values of the same ten sub portfolios. The average deviation was 5%. The figure can only convey an impression of the potential of indicator methods. Neither can it live up to the precision that can be reached by substituting the binary opposition variable for a cardinal value information, nor does it reflect potentially synergistic effects between patents in a portfolio. Elaborations on the imperfections must be left to future studies.

The empirical results are worth of further investigation in the future. As laid out in detail in the discussion of the multivariate results, some limitations of the research design used in this study could be avoided in future work. Ambiguities in the interpretation of individual indicators will diminish if the actual patent value is used as the dependent variable. Corrections for heteroscedasticity should become obsolete if the sample consists of patents from only one company. As mentioned, however, surveys of the last sort are very costly and time consuming and require justification by preliminary results as shown in this study.

Finally, I see an alternative chance to extend on this work in the future that does not require costly primary data as suggested in the previous paragraph. Structural models of patent litigation using information on opposition outcomes appear more appropriate to validate indicators of patent value unambiguously than the research design in this study (See Reitzig, 2003b). But, then again, we must learn to walk before we run.

Acknowledgements

I would like to thank Ralph Schories and Robert Schluetken for supporting my data collection efforts and Emma Aisbett and Aman Bandari for carefully reading the manuscript. The usual disclaimer applies.

References

Albert, M.B., Avery, D., Narin, F., McAllister, P., 1991. Direct validation of citation counts as indicators of industrially important patents. Research Policy 20, 251–259.

Cockburn, I., Griliches, Z., 1988. Industry effects and appropriability measures in the stock market's valuation of R&D and patents. American Economic Review 78 (2), 419–432.

Cohen, W.M., Nelson, R.R., Walsh, J.P., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or not). NBER, Cambridge, MA.

Conolly, R., Hirschey, M., 1988. Market value and patents: a Bayesian approach. Economic Letters 27 (1), 83–87.

EPO, 1998, Annual Report. European Patent Office, Munich.

Grefermann, K., Oppenländer, K.H., Peffgen, E., Röthlingshöfer, K.Ch., Scholz, L., 1974. Patentwesen und technischer Fortschritt. Teil I, Göttingen.

Griliches, Z., 1981. Market value, R&D, and patents. Economic Letters 7, 183–187.

Guellec, D., van Pottelsberghe de la Potterie, B., 2000. Applications, grants and the value of patents. Economic Letters 69 (1), 109–114.

Hall, B.H., Ham-Ziedonis, R., 2001. The determinants of patenting in the US semiconductor industry, 1980–1994. RAND Journal of Economics 32, 101–128.

Hall, B.H., Jaffe, A., Trajtenberg, M., 2000. Market Value and Patent Citations: A First Look. NBER, Cambridge, MA.

Harhoff, D., Scherer, F., Vopel, K., 2003. Citations, family size, opposition and the value of patent rights. Research Policy 32 (8), 1343–1363.

Harhoff, D., Reitzig, M., 2002. Determinants of Opposition Against EPO Patent Grants: The Case of Pharmaceuticals and Biotechnology, CEPR WP. 3645, London, UK. International Journal of Industrial Organization, in press.

Harvey, A., 1976. Estimating regression models with multiplicative heteroscedasticity. Econometrica 44, 461–465.

Lanjouw, J.O., Lerner, J., 1997. The Enforcement of Intellectual Property Rights. NBER, Boston, MA.

Lanjouw, J.O., Schankerman, M., 1999. The Quality of Ideas: Measuring Innovation with Multiple Indicators. NBER, Boston, MA.

Lanjouw, J.O., Schankerman, M., 2001. Characteristics of patent litigation: a window on competition. RAND Journal of Economics 32 (1), 129–151.

Lechner, M., 1991. Testing logit models in practise. Journal of Empirical Economics 16, 177–198.

Lerner, J., 1994. The importance of patent scope: an empirical analysis. RAND Journal of Economics 25 (2), 319–333.

Megna, P., Klock, M., 1993. The impact of intangible capital on Tobins Q in the semiconductor industry. American Economic Review 83, 265–269.

Meyer, M., 2000. Does science push technology? Patents citing scientific literature. Research Policy 29, 409–434.

Narin, F., Hamilton, K., Olivastro, D., 1997. The increasing linkage between US technology and public science. Research Policy 26 (3), 317–330.

Pakes, A., 1986. Patents as options: some estimates of the value of holding European Patent stocks. Econometrica 54 (4), 755–784.

Putnam, J., 1996. The Value of International Patent Rights. Yale University Press, Connecticut, NJ.

Rahn, G., 1994. Patenstrategien japanischer Unternehmen. Gewerblicher Rechtsschutz und Urheberrecht (International) 5, 377–382.

Ramb, F., and Reitzig, M., 2004. Who Do You Trust When Bubbles Burst? A Comparative Analysis of the Explanatory Power of Balance Sheet and Patent Information for the Market Values of German Corporations. Frankfurt/Copenhagen: Discussion Paper.

Reitzig, M., 2002a. Valuing patents and patent portfolios from a corporate perspective—theoretical considerations, applied needs, and future challenges, UNECE Expert Background Paper, OPA/CONF.1/2002/4.

Reitzig, M., 2002b, Die Bewertung von Patentrechten-eine theoretische und empirische Analyse aus Unternehmenssicht. DUV Verlag, Wiesbaden.

Reitzig, M., 2003a. What determines patent value—insights from the semiconductor industry. Research Policy 32 (1), 13–26.

Reitzig, M. 2003b. What do patent indicators really measure? A structural test of 'Novelty' and 'Inventive Step' as Determinants of patent profitability. Copenhagen, LEFIC Working paper 2003–1.

Rivette, K., Kline, D., 2000. Discovering new value in intellectual property. Harvard Business Review 1, 54–66.

Schmoch, U., Grupp, H., Mannsbart, W., Schwitalla, B., 1988. Technikprognosen mit Patentindikatoren. Verlag TÜV-Rheinland, Köln.

Shane, S., 2001. Technological opportunities and new firm creation. Management Science 47 (2), 205–220.

Schmoch, U., Kirsch, N., 1994. Analysis of International Patent Flows. Fraunhofer Institut für Innovationsforschung und Systemtechnik, Karlsruhe, Germany.

Tong, X., Frame, J.D., 1992. Measuring national technological performance with patent claims data. Research Policy 23, 133–141.

Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations. RAND Journal of Economics 21 (1), 172–187.