What Do Patent Indicators Really Measure?
–
A Structural Test of ‘Novelty’ and ‘Inventive Step’ as Determinants of Patent Profitability

This version: December 2003

Markus Reitzig [1]

1 Department of Industrial Economics and Strategy, and LEFIC Center for Law and Economics. The Copenhagen Business School, Solbjergvej 3, 2000 Copenhagen-Frederiksberg/Denmark. Reitzig@cbs.dk. Phone +45 38 15 25 35. Fax +45 38 15 25 40.

What Do Patent Indicators Really Measure?
–
A Structural Test of ‘Novelty’ and ‘Inventive Step’ as Determinants of Patent Profitability

Abstract:

As of today, patent based indicators such as citations are widely used to assess innovative output. Despite the large variety of empirical studies in the field, however, the exact relation between indicators and innovation value is still based on multifarious assumptions that are not unambiguous. This paper provides the first empirical test of patent indicators as value measures in the structural form. At the same time and also originally, the paper empirically tests the fundamental hypothesis that patentability requirements such as novelty and inventive step are positively correlated with innovation value. The study draws from a newly assembled data set comprising the entity of European polymer patents between 1978 and 1990. The estimations are carried out using an original two stage discrete choice model that disentangles effects of technical and other value driving properties of innovations. The results support the assumptions that novelty and inventive step enhance a patent’s value. They confirm the importance of backward citations, family size, and forward citations as value indicators. However, they expand on and partly break with older explanations of why patent indicators correlate with profitability.

Keywords: Patents, opposition/litigation, patent indicators, discrete choice

JEL-Classifications: C25, C51, K41, L00, L20

1. Introduction

In 1965, Frederic Scherer drew economists' attention to the importance of patent data as output measures of industrial innovation. In the ensuing 25 years, an unprecedented stream of research evolved that used patent information as economic indicators. The main results of these two and a half decades of empirical patent economics – mainly studies on industrial productivity – are summarized in Griliches' (1990) widely quoted survey article. Since then, however, activities in the field have not stalled. On the contrary, they have branched out in various directions and the use of patent information has entered into diverse economics and management disciplines. Inspired by the rising electronic availability of patent data and increasing processor speed and memory of personal computers during the last fifteen years, economists have spent extensive time on developing more sophisticated patent indicators than simple patent counts. Major efforts concentrated on the compilation and interpretation of procedural legal information published together with the disclosure of the technical invention underlying the patent. Nowadays, backward citations, forward citations, family size, and claims (to mention but a few) are standard indicators used to qualify patents and weight patent counts. The application of these indicators is no longer restricted to research questions associated with the original difficulty of measuring innovative output (Griliches, 1981; Conolly et al., 1986; Conolly and Hirschey, 1988; Megna and Klock, 1993; Cockburn and Griliches, 1988; Hall et al., 2000). Economists have also applied patent indicators to determine the likelihood of litigation (Lanjouw, 1998; Lanjouw and Schankerman, 2001; Harhoff and Reitzig, 2002), to study patterns of industrial organization (Bekkers al., 2002), and technology spillovers (Verspaargen and Loo, 1999). Management scholars use patent indicators to identify lucrative market segments (Ernst, 1996), for competition analysis (Ernst, 1998), and most recently even to study knowledge flows within corporations (Trajtenberg et al., 1997; Rosenkopf and Nerkar, 2001; Argyres and Silverman, 2002). Looking at this rapid development it seems as if the applicability of patent information to the measurement of economic phenomena was straightforward.

This impression, however, may prove to be deceptive at second glance. This is because the information content of patent indicators is complex and the diversity of its potential meanings is far from being understood in all detail. The main caveat is that the information contained within patents (and respective indicators) is legal of nature and therefore first of all operationalizes (latent) legal variables. Only through an additional body of theory can these indicators ultimately be linked to economic phenomena. During the last 30 years, patent economists have drawn a complex picture of assumptions concerning which *latent variables* drive a patent's economic value both from a welfare and an individual perspective and which trade-offs they are subject to (Scotchmer and Green, 1990, 1995; Scotchmer, 1996; Gallini, 1992, Klemperer, 1990; Gilbert and Shapiro, 1990). Our theoretical understanding of the correlations between a patent's *observable* legal characteristics and their economic effects, however, is still very limited in part. By saying so I am not only referring to the most recent applications of patent indicators to knowledge management, for which little or no hard theory-based empirical evidence exists that patent measures operationalize the latent constructs they are supposed to operationalize. Even in the 'classical' field of indicator-based assessments of intellectual property rights we still often rely on various plausible assumptions and widely accepted connotations when interpreting estimation results. It is indisputable that a large variety of empirical studies in this field (see Appendix A) have given rise to the assumption that procedural patent data are generally suited to operationalize a patent's economic value. To the best of my knowledge, however, no empirical study exists that allows for the interpretation of coefficients of patent indicators as patent value correlates in the structural form (with the exception of the pioneering work on simple patent counts by Pakes, 1986). From a scientific and an applied perspective, however, this is dissatisfying for one major reason: for a variety of theoretical and practical problems we are not only interested in knowing *if* an invention is of commercial value but also *why* it is of commercial value. Industrial economists, for example, often would like to know if sector performances can be attributed to the quality of the underlying technology or to something else. Policy makers need to understand how the adjustment of patentability requirements affects innovation incentives all other variables being equal. Finally, management scholars are interested in the potential of markets independently of the technical value of individual patent rights.

This paper addresses the problem of validating indicators of patent value in the structural form for the first time in the literature. By doing so, it seeks to disentangle the multitude of effects reflected in patent indicators and explain whether and why they operationalize the economic value of patent rights. At the same time, it serves as the first large-scale empirical test of our current theoretical understanding of how patentability requirements affect innovation incentives. As a further product it presents a novel discrete choice estimator suitable for testing decision problems in which the anticipated outcome in the second stage affects the decision in the first stage. The remainder of the paper is structured as follows. Section two provides the theoretical framework and section three presents the hypotheses, the research design, and the underlying econometrics. Part four presents the data and empirical estimation results that are discussed in part five. Section six concludes.

2. Theoretical Framework

During the last two decades, at least two different research directions related to the valuation of patents have developed in parallel. Even though they have occasionally been linked, no systematic approach has yet been chosen to bring the two together until very recently (Reitzig, 2002a). These two research streams are described in the (mainly) theoretical literature on the optimal design of patent systems and the (dominantly) empirical literature on patent indicators.

Latent variables of patent value

The theoretical literature has assumed that the following (latent) variables should affect a patent's value for his/her owner: *Patent duration, novelty and inventive step (non-obviousness), breadth, disclosure, difficulty in inventing around, and dependence on complementary assets*.[2] Since at least the first three of them fall into the category of legal patentability requirements, most of the respective literature in the field stems from economists interested in the design of innovation systems. Already Nordhaus (1967) took his point of departure in the premise that the economic value of a patent for its holder increases with the patent's duration. More recent models (see for example Matutes et al., 1996) differ from their predecessors mainly in that they make more realistic assumptions as to the distribution of returns-per-period over time.[3] Green and Scotchmer (1995) introduced the impact of 'novelty' and 'non-obviousness' (inventive step) on patent value to the discussion. They assumed both variables to increase the value of the patent right. Klemperer (1990) and Gilbert and Shapiro (1990) were the first to propose that the degree to which a patent protects an invention, namely the patent's breadth, affects the patent's value positively. It was again Green and Scotchmer (1995) who modeled disclosure as a value driver of patents assuming that disclosing technical information conferred a positive externality on the patent-holder's competitors. Consequently they assumed that disclosure should diminish the economic value of a patent for his/her owner. Also, as patents may be used for blocking competitors in certain industries, their values should rise the more difficult it becomes to circumnavigate the protected invention with a new technology. Gallini (1992) introduced this idea into a formal model for the first time. Finally, it was Teece (1986) who reminded us that oftentimes,

2 See Reitzig (2003) for an overview.

3 Consistent with the literature on technology cycles (see for example Kotler and Bliemel, 1995) these models do not assume that returns-per-period are constant but that they are subject to the life stage of the underlying technology.

complementary technology and other complementary assets are needed to commercialize the patent protected invention.

More recently, these *latent* variables have been referred to in the management literature as 'value drivers' of the underlying technology in real-option frameworks (Reitzig, 2002b/c) that are currently extended to patent valuations (Pitkethly, 1999). Within these frameworks, the influence of technical (Gilbert and Newberry, 1982), market (Gilbert and Newberry, 1982), and legal (Lanjouw, 1998; Harhoff and Reitzig, 2002) uncertainty on a patent's present value is discussed as the patent's volatility. Whereas the first two types of uncertainty should theoretically enhance a patent's value for his/her owner, the latter one can only reduce it.

Very little direct empirical evidence from *primary data* exists on the validity of the above assumptions. The only existing questionnaire-based study in the field stems from the semiconductor industry (Reitzig, 2003). Its small-scale findings are consistent with the theoretical assumptions apart from the effect of disclosure.

Indicators of patent value

Comprehensive empirical studies using *secondary data* have been carried out trying to validate procedural legal information referring to the granting procedure or to litigation details as value indicators. Appendix A provides a synopsis of the most important large-scale empirical studies ordered by indicators.

The meaning of these indicators as well as their degree of theoretical and empirical validity has been extensively discussed in the literature (see Reitzig, 2002b/c for a comprehensive overview).[4]

---

[4] Readers are referred to the relevant sources for the detailed discussion of the current knowledge on value indicators. For the purpose of this paper, only the most important findings relevant for the derivation of the hypotheses are summarised briefly.

Backward citations – these are quotations of prior art relevant to the patentability of an invention during the granting procedure – are supposed to be positively correlated with value. In the case of quotations to prior *patent* literature one rationale is that the citations operationalize existing market potential. In the case of quotations to the *non-patent* literature – especially scientific publications – economists often argue that the link of the patent's invention to basic research indicates high technological quality and therefore economic value (Carpenter et al., 1980). Backward citations to both patent and non-patent literature belong to the fairly well-validated indicators.

Forward citations, that is the number of quotations a patent receives itself during subsequent granting procedures of younger patents, turn out to be positively correlated with a patent's value in all known studies. Various rationales are put forth. One is that forward citations are supposed to operationalize market potential of the patent independent of the technological sophistication or quality of the underlying invention. This rationale holds especially – but not exclusively – true for citations made by an applicant. At the same time, however, forward citations are also suspected to operationalize the *legal value drivers* novelty and inventive step. This appears particularly plausible if the quotations were inserted by the patent examiner.

'Family size' is an indicator that measures the size of the territory in which the patent holder enjoys exclusivity. Most times it is argued that it is a measure of the market size of the invention which is not necessarily correlated with technical sophistication. Other authors have put forth, however, that family size might be an indicator of technical sophistication as well. The family size indicator is also fairly well validated.

A series of other indicators have been tested as value indicators in earlier studies, too. Among those are the 'Scope' variable, the ownership variable, litigation indicators, indicators referring to the filing strategy, the number of applicants, the number of cross-border research co-operations, the accelerated examination request, and the claims. The empirical evidence of their validity as of today varies; however, they still all belong to the extended set of weighting measures for patent counts applied in the field today. More interesting for the purpose of this paper are the differences in their theoretical foundation. Whereas claims and scope are supposed to operationalize a patent's breadth, the number of applicants and the amount of cross-border research co-operation should reflect a high degree of technological sophistication: i.e. the technology should be novel and highly inventive. The filing strategy variables and the accelerated examination request are deemed to be mainly signs of market proximity of the invention. Finally, the meaning of the litigation indicator depends on its compilation. Counting unsuccessful oppositions or challenge suits *ex-post* yields a pure value indicator. Counting legal attacks *ex-ante* before the outcome of the opposition procedure or the trial is over has a more complicated meaning that is not discussed here but in the section on the research design.

3. Open research questions, hypotheses, and research design

In the introduction I identified one major research goal in the field, namely to understand *why* inventions are of commercial value by using indicators. One important application of this general goal was an empirical test of the fundamental but untested assumption that patentability requirements such as novelty and inventive step affect the economic value of a patent right. Based on current understanding, the following hypotheses should therefore be tested.

Hypotheses

> H1: Novelty and inventive step (non-obviousness) affect the economic value of a patent-protected invention positively.

H2: Indicators such as backward citations to the non-patent literature, forward citations, the number of inventors, and the number of applicants (from now on referred to as indicator set xb) operationalize novelty and inventive step.

H3: Indicators such as backward citations to the patent literature, forward citations, family size, indicators referring to the filing strategy, and the accelerated examination request (from now on referred to as indicator set zg) are positively correlated with a patent's value. However, they do not operationalize novelty and inventive step.

The following section describes the research design chosen for the test of H1 through H3.

The research design

In principle, two generically different approaches can be chosen to test H1 through H3. One potential research design relates indicators to custom tailored *primary data*, namely expert assessments on the latent variables of patent value. This paper pursues the other approach using *secondary data*. Based on a simplified theoretical decision-making problem it analyzes observable oppositions in the European patent system in the light of their expected outcomes by the parties.

The theoretical decision-making problem

In the European patent system, third parties can attack a patent within nine months after its grant by filing a so-called opposition. The opposition procedure differs slightly from the challenge suit procedure in the US (see Reitzig 2002c for details). For the purpose of this paper, these differences can be neglected though. There are three potential outcomes of an opposition procedure. Either the patent is upheld and remains unchanged (1), or the patent is amended (2), or it is revoked (3).[5] According to Art. 100 of the European Patent Convention

[5] Graham et al. (2002) note that there is a 'forth' outcome category, i.e. the opposition procedure is

(EPC), the ruling on the outcome is made by the European Patent Office based on a (re)assessment of the patentability requirements, the main ones being: novelty, inventive step and commercial applicability. Finally, insufficient disclosure of the invention can lead to a revocation or an amendment of the patent, too. Since an opposition procedure is costly and since the alternative option to an opposition is a settlement agreement, one can assume that

> *A1: The incentives to file an opposition are determined by the value at stake and the likelihood, as subjectively perceived by the parties, of the patent being upheld, amended, or revoked.*

In the economic literature on litigation an analogous assumption has been used extensively (Priest, 1984; Waldvogel, 1998) and appears to be commonly accepted. Lanjouw and Lerner (1997) also showed the suitability of this premise in the case of *patent* litigations. For the specific case of an *opposition* Harhoff and Reitzig (2002) emphasize the particular importance of settlement costs, but in principle they also agree on the above assumption.

Thus, the opponent's rationale to file an opposition can be illustrated by the decision tree in Figure 1.

Insert Figure 1 about here

"closed". The way they define this category it "refers to cases in which the patent holders do not renew patent protection after the opposition has been filed, which causes the patent to lapse into the public domain". For want of better information they suggest considering these cases as successful challenges of the patent's validity. While there may be explanations for doing so, this paper takes a different approach based on intense discussions with a senior representative at the Technical Board of Appeals at the European Patent Office. The expert argues that the 'closure' of an opposition as reported in the data source www.epoline.org is a decision made by the opponent and most likely a 'retreat' from the legal attack.

The (potential) opponent will attack the patent holder if his/her profits in the case of an opposition exceed his/her profits in the case of passive behavior. His/her decision-making problem and the resulting likelihood of an observable opposition can thus be expressed in formal terms.

$$W = 1 \; if$$

$$p_{rj} \cdot p_{opponent}(rejec \mid opp.) + p_{am} \cdot p_{opponent}(amend \mid opp.) + p_{rv} \cdot p_{opponent}(revoc \mid opp.) > p_{no} \qquad (1)$$

$$and$$

$$W = 0 \quad otherwise$$

The estimation problem and its relation to the hypotheses tests

To turn the system of equations (1) into an estimation problem that allows for a test of H1 through H3, some further assumptions and simplifications are necessary. The assumptions are mostly unproblematic. Thus, they are only briefly described and vindicated in the following.

> *A2: The opponent's profits can theoretically be driven by both patentability requirements (novelty, inventive step) and other non-technical factors.*

This assumption is straightforward. Any alternative assumption would falsify H1, H2, and H3 *per definitionem.*

A3: The opponent can reasonably anticipate the decision of the opposition procedure by the European Patent Office (EPO). The EPO bases its ruling on the opposition solely on two criteria, namely the fulfillment of novelty and inventive step.

This assumption is necessary to infer from observable outcomes of opposition decisions (see below) on the opponent's estimation of novelty and inventive step. The second part of the assumption is entirely unproblematic because it reflects the dogmatic guideline of the EPO, but the first part is also very plausible given the large litigation experience of most opponents (see also Harhoff and Reitzig, 2001). In a formalized fashion, this assumption states that

$$p_{opponent}(outcome \mid opp.) = f(novelty, inventive\ step) \qquad (2)$$

Finally, the last two assumptions are the following.

A4: The opponent's profits given the different possible rulings of the EPO are determined as follows:

$$p_{no} = 0; p_{rj} = p_{am} = -c; p_{rv} = g(novelty, inventive\ step, other\ \ factors) - c$$

(with c = costs for an opposition)

and

A5: The opponent's decision to file an opposition is not significantly codetermined by the possibility to appeal against the EPO ruling on the opposition procedure.

Assumption 4 is the most simplifying of all. It states that the opponent almost completely internalizes the value of the patent for his/her holder which might not be entirely true for all competitive scenarios.[6] It also contains a simplification in that is sets the profits for two outcome scenarios equal, namely the amendment of the patent and the rejection of the opposition. From talks with patent attorneys (Reitzig 2002c), however, it seems that this simplification also reflected reality sufficiently well in certain industries (see also section “Empirical Results – sample selection”). Finally, Assumption 5 potentially simplifies the opponent’s rationale in that it no longer considers subsequent appeal or litigation possibilities in detail.[7]

Implementing A1, A2 and A4 into (1) then yields the following condition for an opposition:

$$
\begin{aligned}
&\textit{With}\quad p = p(\textit{amendment} \mid \textit{opp.}) + p(\textit{rejection} \mid \textit{opp.})\\
&W = 1 \; \textit{if} \;\; p(p_{am,rj} - p_{rv}) > (p_{no} - p_{rv})\\
&\textit{and}\\
&W = 0 \quad \textit{otherwise}
\end{aligned}
\qquad (3)
$$

To test H1 through H3, a maximum likelihood estimator based on Equation (3) is now needed. Using a corresponding ML-estimator and operationalizing novelty and inventive step by indicator set xb (and error term e) and other factors by indicator set zg (and error term h) could disentangle the multitude of effects measured by indicators. It could finally help to explain *why* patent indicators measure patent value. Despite the great variety of various two-stage discrete choice models described in the econometric literature, however, no

6 In a sector in which patents are mainly used as bargaining chips, this assumption could lead to a distraction from the ‘real’ results. For the industry analyzed in this paper the assumption should hold well though (see also the following section “Empirical Results – sample selection”).

7 The simplification seems inevitable for both theoretical and practical reasons. From a theoretical perspective, a three-stage model accounting for appeal or other litigation possibilities following opposition does not appear feasible, nor does a further truncation of the data that comes with the observation of appeal cases (see section four) appear reasonable. For practical reasons Assumption 5

standard estimator suitable for this decision problem appears to exist. For this reason, a custom-tailored estimator was developed for the present problem.[8] The likelihood function is given below; its derivation is described in Appendix B.

$$\log L = (1-W)\cdot\log\left(\Phi\left(-\frac{a}{s_x}\right)\right)+W\cdot\left[V\cdot\log\left(\Phi^2\left(\frac{a}{s_x}, xb, r_{x,e}\right)\right)+(1-V)\cdot\log\left(\Phi^2\left(\frac{a}{s_x}, -xb, -r_{x,e}\right)\right)\right] \quad (4)$$

, *where*

$W = 1$ *denotes the occurence of an opposition* ($W = 0$ *otherwise*),

$V = 0$ *denotes the revocation of the patent* ($V = 1$ *otherwise*),

$a = (1-\Phi(xb))\cdot(xb + zg) - c,$

$x = (1-\Phi(xb))\cdot(e + h)$, *and*

$$r_{x,e} = \sqrt{\frac{(1 + r_{h,e})}{2}}.$$

The estimator described in Equation (4) maximizes the likelihood of an opposition including the opponent’s anticipated probability of winning or losing his case.[9] The value of the patent is modeled by novelty, inventive step, and other value determining parameters; the opponent’s anticipation of the patent being upheld (revoked) is modeled by novelty and inventive step (consistent with Assumption 3).

---

also appears necessary, however, slightly stronger at the same time (see FN 13 for an elaboration of the practical problems).

8 Full identification was numerically established using simulated data.

9 This feature distinguishes the estimator from known discrete choice models that are available in statistical software packets, such as multinomial/nested logits or Heckman’s probit (see e.g. Heckman, 1979). Neither do econometric textbooks (see e.g. Maddala, 1983) report on an estimator of the above kind that models a selection bias but focuses on the estimation of the unselected (stage 1) rather than the selected sample (stage 2).

Finally, the data prerequisites of the chosen research design and the consequences for the sample selection are briefly elucidated in the last part of this section.

Data prerequisites

The model underlying the estimator is based on the premise that the value of a patent revocation for the opponent is proportional to the value of the patent for its holder. This assumption is best fulfilled in markets where patent holders enjoy temporary monopoly profits and would have to share duopoly profits if they did not have legal protection. Therefore, running the empirical test on patent data from a discrete product industry seems inevitable.

Besides, a main distinguishing feature from earlier empirical studies on patent litigation or opposition is the explicit modeling of the key assumption that anticipated success and failure rates enter the opponent's or plaintiff's rationale.[10] In the present paper, *observable outcomes* of opposition procedures are used to model the opponent's anticipation. Thus, the data need to comprehend the sets of value indicators as described in H2 and H3, the observable opposition, and the corresponding EPO ruling. This means, however, that the data for the analysis are in principal truncated and non-trivial sample selection problems arise that are described in the next section.

4. Empirical results

[10] To the best of my knowledge the study by Graham et al. (2002) is the only one that presents litigation and opposition outcome data at all. That study, however, contains no structural validation of value indicators modeling subjective outcome anticipations in any way.

Sample selection

The sample chosen for the analysis contains European polymer patents for two reasons.[11] First, it seems that in the case of polymers individual patents could protect most of the technology inherent in the final product (see e.g. Cohen et al., 2000). Thus it is plausible to assume that the value of a patent revocation to the opponent is highly correlated with the value of the valid patent for its owner. Secondly, to the best of my knowledge there is no large-scale empirical study on the European polymer industry. Beyond its primary goal of testing value indicators structurally, the sheer data presentation in this paper already extends our knowledge base of this industry in a more general fashion.

The selection of the industrial field was based on an updated version of the widely accepted OST INPI ISI classification by Schmoch (1998, personal note). Polymer patents were identified as showing one of the following IPC subclasses as their main classification: C08B, F, G, H, K, L; C09D, J. As of October 2003 (date of the data extraction) the European patent register contained 31,178 granted patents in these areas. At this point, 3,126 (10.03%) patents in the sample had been opposed. For 2,608 of these opposed patents, a decision by the *first instance* at the EPO – the opposition division – was observable in October 2003. For the remaining part of the patents I could not identify any clear ruling by the opposition division at that date from electronic sources.[12] Figure 2 shows the share of unidentified oppositions among the total sample versus the year of patent priority.

---

11 The patents were identified via the OST INPI ISI classification based on IPCs.

12 The problem of identifying opposition outcomes correctly is very subtle, both for theoretical and practical reasons. In principle, there can be two reasons why an opposition outcome may not be easily computable from an electronic source. Either (1) the case is still pending in the opposition division itself and then there is simply no outcome observable or (2) the case went into appeal but is not yet decided by the Board of Appeals. Why the second case leads to a problem of identifying opposition outcomes requires some further explanation. In publicly accessible databases (such as www.epoline.org), data fields on the information of the first ruling by the opposition division are overwritten by the ruling of a later appeal decision in the case of an appeal. In pending appeals, the data storage procedure of www.epoline.org thus 'feigns' an increased share of undecided first instance opposition outcomes. However, given that this paper concentrates exactly on the first instance opposition outcomes (i.e. the rulings by the opposition division), the European register data are therefore not suited for this analysis without further consolidation. Consequently, the data were consolidated using an additional data source, namely the Board of Appeals Database (to the extent that its decisions are published). In many appeal

Insert Figure 2 about here

Figure 2 shows a steep ascent of unidentified opposition outcomes after the year 1990. Pending better explanation I take it that this increase can be attributed to the share of opposition cases still to be decided in October 2003 by the first instance at the EPO, namely the opposition division.[13] As this paper (see Assumption 5) primarily focuses on the decision of the opposition division (no further appeals, no subsequent litigation), I cut off the tail of patents applied for from 1991 onwards and obtain a residual percentage of approx. 2.1% of unidentified first ruling opposition cases between 1978 and 1990.[14] Thus, the final data for the analysis comprises 17,123 granted EP patents with priority dates between 1978 and 1990, of which 2,034 were opposed.

---

cases, the first instance opposition outcomes can be recalled from the historical case files stored in this database. Thus, in an attempt to 'clean' the dataset for this analysis and obtain correct opposition outcomes, extensive full text cross-checks were carried out on the databases of the EPO's Board of Appeals to reassemble the historical first instance opposition rulings in case of subsequent appeals. After this effort only a very small number of opposition outcomes remain unidentified. The share of those yet undecided opposition outcomes should increase with more recent application, as is confirmed in Figure 2.

Note: Despite this major effort there remain some potential problems associated with the opposition outcome data. For example, a declaration borderline case exists when the Board of Appeals refers the case back to the opposition division. In this case, it is the opposition division that decides the case. However, one may criticize the fact that the decision takes the rationale of the Board of Appeals into account (see Assumption 5). On the other hand, the decision is a formal decision by the opposition division and may well differ from the opinion expressed by the Board of Appeals as it seems to this author (see EP12928 as an example).
Whereas the latter problem is a definitory (hence theoretical) problem, there exists a purely practical problem in that even the most sophisticated automated programming routines cannot cover every special case that may be observed in the data. While the identification routine for opposition outcomes, including the automated full text cross-checks in the appeals' database, should in most cases allow for the correct distinction between the opposition's and the appeal's outcome, residual disturbances cannot be excluded because of the size of the dataset. Given the fact, however, that the "error" resulting from these imperfections only lies in the potential confusion of the *instance* of the decision (opposition vs. appeal) in a few cases (leading to a minor relaxation of Assumption 5) I desisted from a comprehensive manual data compilation (the latter being the only possibility to exclude all remaining uncertainties).

13 Data on the average duration of opposition procedures have been published in the literature for the fields of pharmaceuticals and biotechnology only. As Graham, Hall et al. (2002) describe, opposition procedures may take about 2.7 (2.8) years (post/pre 1991 applications). Adding the average granting time of 4.3 years for patents in the same area (Reitzig, 2002b) to that value yields an average period of 7 years from grant to opposition ruling in these industries. Looking further at the variance of this period it is therefore entirely plausible to observe an increase in undecided opposition procedures of 11 to 12-year-old patent applications as is the case here.

14 The resulting "imperfection" of the data set appears acceptable considering that with 2.1%

Descriptive statistics

Table 1 shows the descriptive statistics for the sample.

Insert Table 1 about here

The most interesting findings are briefly discussed. At 11.87% the rate of opposition in the polymer chemistry industry is significantly larger than in the total population of EP patents of this period. It is even higher than in the litigious pharmaceutical and biotechnological industry (10.79% and 10.24% opposition between 1978 and 1990). This preliminarily confirms the assumption that patents in polymers are used as 'exclusion rights'. In about 39% of all oppositions, third parties attack holders successfully according to the notion of this paper; i.e. the patent is revoked by the opposition division. In about 28% of the cases the patent is amended, in another 26% of the observations the opposition is rejected because it is not considered to be substantiated. Both cases are considered to be defeats for the opponent in the current paper.[15] In the remaining roughly 6% of the oppositions the procedure was either closed (4.7%) or no outcome can yet be identified (see above, 2.1%). The latter two cases are also regarded as defeats for the opponent.[16]

---

unidentified cases the "disturbance" of the opposition outcome variable is negligible.

15 The reason is that in a discrete product industry such as polymer chemistry amendments can theoretically be backed up fairly well by the integration of so-called fall-back options. Those are inserted in the form of dependent claims in the patent draft (Reitzig, 2002b). Thus, it is likely that amendments are less harmful for the patent owner than in other industries.

16 With respect to the closures, this appears plausible for two reasons: Firstly, talks with an expert from the Board of Appeals reveal that this interpretation appears suitable (see above, footnote 6). Secondly, there is no indication that these patents – if lapsed – were lapsed *because* of the opposition procedure. Besides, the share is very small.

With respect to the unidentified cases, the following argument can be put forth to consider them defeats for the opponent. The only strategic value for an opponent to prolong the opposition by an appeal (all other variables being equal) is to create uncertainty on the side of the patentee. Since product cycles are relatively short in the polymer industry compared to pharmaceuticals, however, it is reasonable to assume that for an opponent a successful attack consists in achieving a fast ruling on the invalidity of the patent. Again, however, the share is so small that interpreting undecided cases either way most likely does not change results substantially.

The age structure of the sample reflects the general application trend at the European Patent Office. About 15% of the patents have priority dates lying between 1977[17] and 1980. Roughly 35% of the patents in the sample refer to priorities between 1981 and 1985, and the rest of the patents to priority dates between 1986 and 1990.

The explanatory variables are not conspicuous in that their order of magnitude corresponds to various earlier studies in related industries, in particular to the study by Reitzig (2002a). On average, three and a half references to patents of prior art were made by the EPO examiners during the European search procedure.[18] On average, six out of ten patents cite a non-patent literature reference as relevant state of the art. The patents were applied for in 7.5 states on average, and almost three inventors (2.7) were involved in each application. The relatively high number of designation states and the high inventor-to-applicant ratio supports the assumption that most of the patents in this industry are held by corporations and not by individual inventors. This observation is in accordance with the high opposition rate given that oppositions are costly. The mean for accelerated examination requests is fairly low. Not even one percent of all patents are applied for following the Programme for Accelerated Prosecution of European Patent Applications (PACE). This observation supports the view that lead-time advantages in this industry may be less important than in other industries. It also supports the assumption, however, that the applicants are experienced in interacting with the EPO and can anticipate the office's reactions well (see Assumption 3).[19] Interestingly, however, the percentage of filings according to the Patent Corporation Treaty (PCT) is high compared to Reitzig (2002a). The latter findings indicate that applicants are delaying cost intense decisions in more than 7% of the applications by choosing the PCT route. While this

---

17 Note that while the European Patent Office was not inaugurated until 1978, priority applications may date back until 1977.

18 Note: As Harhoff and Reitzig (2002) and Reitzig (2002a) this paper adds the patent references of the international search to the number of references made in the European search if the patent was a PCT patent and the EPO acted as the International Search Authority for the World Intellectual Property Organization.

19 Acceleration of examination if oftentimes requested if the decision of the EPO cannot be anticipated at a point where investments have to be made (see Reitzig, 2002b).

may be a sign of uncertainty (and therefore be at odds with the low acceleration request rate) an explanation for the observation lies in the applicant structure of this sample. Given that the applicant-to-inventor ratio hints at a dominating corporate applicant structure it may well be that the applicants are simply cost insensitive. Finally, the average number of forward citations (three year's time window after publication) in subsequent EPO search procedures is 0.66[20]. This figure is fairly low compared to Lanjouw and Schankerman (2001) or Harhoff and Reitzig (2002) indicating that the average scientific impact of a polymer patent on subsequent applications may be lower than in pharmaceuticals or biotechnology. However, these considerations remain speculative to a certain extent. Thus, in the following the structural estimation seeks to contribute to a somewhat better understanding of what the indicators really measure. Before that, however, reduced form estimation results are presented.

Reduced Form Estimates

Table 2 shows the results of four simple probit estimations. In column A the probability of an opposition is modeled by the set of indicators based on the entire sample. In column B the conditional probability of the patent being upheld after opposition is modeled on the same set of indicators. However, regression B is carried out on the sample of opposed patents only. Given the fact that the sample spans over a time period of 13 years it seems reasonable to control for pure time effects. Hence, columns 2A and 2B show two different estimations each, one without and one including time dummies.

Insert Table 2 about here

20 This figure was again corrected for PCT filings where results from the international search procedure of the EPO in its function as the World Intellectual Property Organization's (WIPO) international search authority must be taken into account.

Overall, the models of the probability of an opposition presented in column A are well determined as can be seen from Table 2. In specification 2.A.I. individually significant coefficients are found for the backward citations to the patent literature and to the non-patent literature, the family size, and the forward citations. The remaining variables are individually and jointly insignificant ($\chi^2(4)$-test: 0.99; P>0.91). Introducing time dummies for the priority periods between 1980 and 1985 and between 1986 and 1990 (reference category priority period 1978 until 1984) in specification 2.A.II. renders the coefficients for the number of inventors and the accelerated examination request significant, too, without changing the levels of significance for those coefficients that were already significant in 2.A.I. Looking at the first specification in column B, the family size, the number of inventors, and the forward citations are significantly and positively correlated with the maintenance of the patent in an opposition procedure. The remaining variables are individually and jointly insignificant ($\chi^2(5)$-test: 3.39; P>0.63). However, the overall model is not well determined. This changes, however, when time dummies are introduced. Interestingly, the family size indicator also loses its significance when time dummies are used as control variables (see 2.B.II.). Correlations among the independent variables are moderate overall (-0.10 to 0.23) and are therefore not reported separately here. The only stronger correlation (0.42) is found between the PCT indicator and the number of backward citations to the non-patent literature. The finding is interesting in that it hints at a systematic difference between the non-patent reference search outcomes for PCT filings and pure European filings – a question that is not of central interest for this paper, though.

Turning back to the scope of this article, on the basis of the simple estimation results in column A it is impossible to test H1, H2, and H3. The reason is that it is impossible to distinguish between the *different effects* that the individually significant variables exert on the likelihood of an opposition. Whilst it may be – as has been discussed recently in the literature – that the backward citations to the patent literature hint at the existing market and that the negative coefficient for the non-patent literature references hints at lacking market proximity of the invention, various other explanations may also hold true. It is by no means possible to make a clear statement as to whether novelty and inventive step affect the value of the patent positively or not. Looking at the regression results of column B it would also be difficult to reject or sustain hypothesis H2 in a general fashion. Undoubtedly, the results are very stimulating in that they show that for the subset of opposed patents correlations exist between a set of indicators and the likelihood of the patent being upheld. However, the estimations carried out in column B are carried out on a subset of patents that entered an opposition procedure either because of their high value for the opponent or because of their low legal and technical quality. Thus, the coefficients in column B are biased and may potentially operationalize various effects, including effects that are unrelated to the degree of novelty and inventive step. Thus, to test H1, H2, and H3 only the estimations in the structural form can shed further light on the research questions.

Estimations in the structural form

Table 3 shows the results of the new simultaneous equation model that reflects the decision-making rationale of the opponent in stage one (opposition) depending on his anticipated outcome in stage two (EPO ruling on the opposition outcome)

Insert Table 3 about here

According to the above notation, XB denotes the set of indicators that are supposed to operationalize novelty and inventive step. Correspondingly, ZG denotes the set of indicators that are supposed to be value correlates of the patent without operationalizing novelty and inventive step.

To test H1 through H3, specification 3A starts from the (partially) disjunct set of available explanatory variables assumed to operationalize the different value drivers of the patent without controlling for time effects. The major result is partly encouraging and partly discouraging at first sight; the model is significant overall (χ^2(9)-test: 21.28; P=0.01), though on the lower end of the spectrum. Thus, from a purely statistical standpoint it models the likelihood of an opposition worse than the simple probits in column 2A. As was argued above, however, this simple statistical measure does not reflect the economic suitability of the results. Model 3A, though weakly defined overall, hints at the importance of two indicators of novelty and inventive step, namely the number of backward citations to the non-patent literature and the forward citations that are both individually significant variables in XB. While encouraging in general terms, the result will have to be discussed in more detail in the following, as the negative coefficient for the references to the non-patent literature requires further discussion. The remaining variables in XB are individually and jointly insignificant (χ^2(2)-test: 0.05; P>0.96). Similarly, the results for the coefficients in ZG are also interesting. Without over-interpreting specification 3A, it is interesting that the individually significant coefficients for the patent backward citations, the family size, and the forward citations are consistent with theoretical expectations (H3). Again, however, one of the coefficient's signs requires further discussion, namely the coefficient for the forward citations. The remaining variables in ZG are individually and jointly insignificant (χ^2(3)-test: 2.44; P>0.48). Finally, the negative sign for s_h lacks an eidetic meaning.

To analyze whether some of the counterintuitive findings can be exposed as artifacts that stem from underlying time effects in the sample, specification 3B includes the time dummies for the different periods as described above. The first thing to note is that specification 3B is well defined overall (χ^2(13)-test: 102.41; P<0.01) as can be seen from Table 3. The finding is elating in that it illustrates the explanatory power of the chosen structural estimation in this paper much more clearly than specification 3A. Besides, the coefficient for s_h is very plausible. With respect to the individual variables, various observations can be made. The introduction of the time dummies leads to an increase of the significance of the number of inventors in XB from 3A to 3B, however, the coefficient for the number of citations to the non-patent literature remains significantly negative. Interestingly, however, the counterintuitive finding for the number of forward citations in ZG vanishes as the coefficient becomes insignificant. Simultaneously, and consistent with the expectation, the dummy for the accelerated examination request becomes significant.

5. Discussion

The results from the multivariate estimations require further interpretation before conclusions can be drawn with respect to the validity of the hypotheses H1 through H3. To do so, it seems reasonable to concentrate on the potential meaning of the structural regression results. However, they are contrasted with reduced form estimation results whenever counterintuitive findings need to be explained.

In principle, the fact that in model 3B two of the individually significant variables in XB, namely the number of forward citations and the number of inventors, contribute positively to the likelihood of an opposition can be regarded as strong evidence to support hypothesis H1. The reason is that the estimator uses the indicators in XB simultaneously for both estimating the likelihood of an opposition and the likelihood of patent maintenance. Thus, variables with significant and positive coefficients in XB operationalize novelty and inventive step and contribute positively to the patent's value.[21] The negative and significant coefficient for the number of citations to the non-patent literature (also found in the reduced form estimations) cannot contradict the overall finding; however, it reduces its clarity. Looking at the effect sizes of the individually significant coefficients in 3B and their

[21] Note that theoretically the result can still be an artifact if the indicators operationalized further

variables' means in Table 1, the overall effect on the likelihood of an opposition exerted by the entity of the variables in XB should be positive. Hence, specification 3B supports H1. This finding appears to be of considerable relevance with respect to the large amount of theoretical literature in the field designing patent systems on the assumption that H1 holds true.

Following the same line of thought I therefore also find strong empirical evidence that forward citations made by the examiner as well as the number of inventors operationalize novelty and inventive step as was hypothesized in H2. As mentioned before, the significant negative coefficient for the number of backward citations to the non-patent literature in XB is unexpected. Breaking with the reduced form estimates of 2.B.I./II., purely technically speaking the structural estimation finding suggests that non-patent literature references are an indicator of lacking novelty and inventive step – a finding that is at odds with those streams of literature that propose the value-indicating properties of this variable because of science linkage reflection. One explanation for this finding that could not have been seen from the simple reduced form estimations may be that the current paper did not distinguish between 'scientific' non-patent literature references and other non-patent literature references. As Meyer (2000) notes, there exist various types of non-patent citations, and not all of them reflect references to basic research. Moreover, Meyer (2000) finds that citation linkages to non-patent references often hardly represent a direct link between a cited paper and a citing patent. Ultimately, it is up to further tests to shed more light on the partly counterintuitive finding of this paper. For reasons of feasibility those tests must be left for future research, though.

Finally, the paper provides evidence for the validity of H3, however, not without teaching a more differentiated story than was expected. All of the individually significant indicators in ZG of specification 3B, namely the number of citations to the patent literature,

---

contingency variables not inherent in the model. There is no reason to believe that this is the case the family size, and the accelerated examination request, show the expected signs. Interestingly, however, only the estimations in the structural form support H3. Whereas the reduced form estimations 2.A.I. & II. could not support H3 due their ambiguity problem, specification 3B unambiguously shows that non-technical aspects of patents contribute to their economic value[22], and that those non-technical aspects are in part reflected in some of the indicators the patent system generates itself. At the same time the paper breaks with the common notion that forward citations would reflect technical as well as technology unrelated aspects of a patent's value. The current findings suggest that forward citations are correlated with a patent's value because they operationalize its technical importance.[23] On the other hand, backward citations to the patent literature or the family size operationalize a patent's value in that they are correlated with non-technical economic features of the property right. The latter finding supports earlier suspicions by Harhoff and Reitzig (2002) that backward citations to the patent literature are an indicator of market size.

Finally, a brief look shall be taken at the individually insignificant coefficients in estimation 3B. The number of applicants – hypothesized to indicate a patent's value for technical reasons – is insignificant. The results are aligned with the reduced form estimates. It appears as if, at least for the existing sample, the impact of 'joint R&D ventures' on patent value was irrelevant. The same irrelevance must be certified for the PCT indicator. Like Harhoff and Reitzig (2002), this paper finds no empirical evidence for the value-indicating properties of the PCT indicator when tested in a joint specification with forward citations and when taking citations of the EPO in its function as the World Intellectual Property Organization's (WIPO) international search authority into account. Whereas the PCT indicator loses some of its value for those studies where forward citations and international

---

though.

22 The term "contribute" does not insinuate a causal relationship that cannot be tested here.

23 See FN 22.

search backward citations can be computed easily, its function to serve as a substitute indicator for ‘young’ patent portfolios (see Reitzig, 2002a) remains unaffected.

6. Summary and future research

This paper began with the premise that various intuitions and assumptions exist as to what determines the value of a patent and how the value of a patent can be assessed using indicators. Despite the existing large-scale empirical evidence in the field it argued that tests of these assumptions still deserve further attention by empirical economists. It proposed that an answer to the question of *why* inventions generate profits at all is crucial for various reasons, namely to understand how patentability requirements related to novelty and inventive step affect innovation incentives and how non-technical factors drive profitability. Ultimately, the question as to how these different *latent* value drivers can be assessed through indicators is of importance to empirical industrial economists. To enhance our understanding, this paper presented an empirical study based on 17,123 EP polymer patents granted between 1978 and 1990. Applying a novel two-stage discrete choice simultaneous equation estimator to the data I could – within my framework of assumptions – support the hypothesis that novelty and inventive step contribute positively to a patent’s value from an individual perspective. Contrasting the structural estimation results to regressions carried out in the reduced form I could maintain the common notion that forward citations and the number of inventors working on one patentable invention are a measure of the patent’s techno-economic quality. I could support the hypothesis that backward citations to the patent literature are correlated with the patent’s value, however, without necessarily operationalizing the patent’s technical sophistication. Most likely, they operationalize an existing market and hint at an increased profitability for that reason. Similarly, the accelerated examination request and the family size are correlated with a patent’s value, however, without necessarily operationalizing novelty or inventive step. However, I also found that the backward citations to the non-patent literature are negatively correlated with a patent's novelty and inventive step – at least as long as they are not distinguished between scientific and other references.

The results are relevant in that they sustain the notion found in the theoretical literature on the design of patent systems that patentability requirements such as novelty and inventive step can be important setscrews for innovation incentives in certain industries. In the sample of polymer patents technical sophistication as operationalized by forward citations seems to be the crucial element for patent value – a result that was not necessarily to be expected and may well be falsified in other industries. On the other hand, patents that over-proportionally 'build on' non-patent references lose in value. In combination, the latter two findings suggest that technical sophistication leads to valuable property rights, however, that this is not necessarily reflected in the number of non-patent references. The results also enhance empirical economists' understanding as to what they really measure when using patent indicators. Finally, the paper contributed a new estimator to the field of empirical economics that shows a special feature in that it simultaneously estimates a discrete choice problem in the first stage conditional on the anticipated outcome of the decision in the second stage. This type of estimator may be useful for various analogous applications as well.

At this point I see various directions in which to proceed with this research. I only want to mention the two most important ones in my eyes. First of all, shortcomings in a current design always open up paths for improvements. In this particular case, more sophisticated distinctions between self-citations and other citations as well as distinctions between type A, X, and Y citations (see Harhoff and Reitzig, 2002; Reitzig, 2002a/b) may be points for potential melioration. In light of the enormous efforts spent on compiling the current data set up to this point, this exercise must be left to future research. Given the rather marginal alterations to be anticipated, however, I also foresee more interesting elaborations along other dimensions. For example, the estimator can be applied to various other industries so that inter-industry comparisons become possible. It is very likely that these comparisons will supply us with interesting novel insights as to how innovation is driven in different industries. It might well be that the existing decision-making model, which is based on the assumption that patents are used as exclusion rights, will fail in other technical fields such as semiconductors and need adjustment. Moreover, relaxations of the assumptions by refining the model structure would be worthwhile. For example, the aggregation of different outcomes in opposition procedures to two generic categories, namely success or defeat, can be defined and might enhance the estimation quality. To do so, suitable proxies for the breadth of a patent would have to be derived, too. Finally, an obvious extension of this research is to apply a somewhat modified estimator to a similar problem in empirical patent economics: the granting procedure.

Acknowledgements

I gratefully acknowledge the invaluable support by François Laisney (ZEW Mannheim and Université Louis Pasteur Strasbourg) and Bronwyn Hall (UC Berkeley and NBER) in setting up the likelihood of the model. Thanks also go to Ralph Schories (IPECON) for extracting the raw data and assembling the indicators. Finally, I thank the Danish Patent and Trademark Office (DKPTO) for its financial support during my time as a faculty member at the Copenhagen Business School. The usual disclaimer applies.

References

Abramson, H. N., J. Encarnacao, P. P. Reid and U. Schmoch (1997). Technology Transfer Systems in the United States and Germany. Washington, National Academy Press.

Albert, M. B., D. Avery, F. Narin and P. McAllister (1991). "Direct Validation of Citation Counts as Indicators of Industrially Important Patents." Research Policy **20**: 251-259.

Argyres, N., and B. Silverman (2002). R&D, Organization Structure, and the Development of Corporate Technological Knowledge, Boston University and University of Toronto.

Carpenter, M., M. Cooper and F. Narin (1980). "Linkage Between Basic Research Literature and Patents." Research Management **(March)**: 30-35.

Cockburn, I. and Z. Griliches (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., R. R. Nelson and J. P. Walsh (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or not). NBER WP 7552.

Conolly, R., B. Hirsch and M. Hirschey (1986). "Union Rent Seeking, Intangible Capital, und Market Value of the Firm." Review of Economics and Statistics **68**(4).

Conolly, R. and M. Hirschey (1988). "Market Value and Patents: A Bayesian Approach." Economics Letters **27**(1): 83-87.

Bekkers, R., G. Duysters, B. Verspagen. (2002). Intellectual property rights, strategic technology agreements and market structure The case of GSM, Research Policy **31**(7): 1141-1161.

Ernst, H. (1998). Patent portfolios for strategic R&D planning. Journal of Engineering and Technology Management **15**: 279-308.

Ernst, H. (1996). Patentinformationen für die strategische Planung von Forschung and Entwicklung, Wiesbaden, DUV-Verlag.

Ernst, H., C. Leptien and J. Vitt (2000). "Inventors Are Not Alike: The Distribution of Patenting Output Among Industrial R&D Personnel." IEEE Transactions of Engineering Management **47**(2): 184-199.

Gallini, N. T. (1992). "Patent Policy and Costly Imitation." RAND Journal of Economics **23**(1): 52-63.

Gilbert, R. and D. Newberry (1982). "Preemptive Patenting and the Persistence of Monopoly." American Economic Review **72**(3): 514-526.

Gilbert, R. and C. Shapiro (1990). "Optimal Patent Length and Breadth." RAND Journal of Economics **21**(1): 106-112.

Graham, S.J.H., B.H. Hall, D. Harhoff, D.C. Mowery. (2002). Post-Issue Patent “Quality Control”: A Comparative Study of US Patent Re-examinations and European Oppositions, NBER WP 8807.

Green, J. R. and S. Scotchmer (1995). "On the Division of Profit in Sequential Innovtion." RAND Journal of Economics **26**(1): 20-33.

Griliches, Z. (1981). "Market Value, R&D, and Patentes." Economic Letters **7**: 183-187.

Guellec, D. and B. van Pottelsberghe de la Potterie (2000). Analysing Patent Grants. Economics Letters, **69**(1): 109-114.

Hall, B.H. and R. Ham-Ziedonis (2001). "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995," RAND Journal of Economics **32**: 101-128.

Hall, B. H., A. Jaffe and M. Trajtenberg (2000). Market Value and Patent Citations: A First Look, NBER WP 7741.

Harhoff, D. and M. Reitzig (2001). "Strategien zur Gewinnmaximierung bei der Anmeldung von Patenten: Wirtschaftliche and rechtliche Entscheidungsgrößen beim Schutz von Erfindungen." Zeitschrift für Betriebswirtschaft **5**: 509-530.

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

Harhoff, D., F. Scherer and K. Vopel (1999). Citations, Familiy Size, Opposition and the Value of Patent Rights. Munich, Boston, Mannheim, Ludwig-Maximilians-Universität München, Harvard University, ZEW Mannheim.

Heckman, J. (1979). Sample Selection as a Specification Error, Econometrica 47, 153-161.

Klemperer, P. (1990). "How Broad Should the Scope of Patent Protection Be?" RAND Journal of Economics **21**(1): 113-130.

Kotler, P. and F. Bliemel (1995). "Marketingstrategien für die Phasen im Produkt-Lebenszyklus." Marketing Management, 557-596.

Lanjouw, J. O. (1998). "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value." Review of Economic Studies **65**: 671-710.

Lanjouw, J. O., A., and J. Putnam (1998) "How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data," The Journal of Industrial Economics. **XLVI** (4): 405-33.

Lanjouw, J. O. and J. Lerner (1997). The Enforcement of Intellectual Property Rights. Boston, NBER WP 6296.

Lanjouw, J. O. and M. Schankerman (1999). The Quality of Ideas: Measuring Innovation With Multiple Indicators, NBER WP 7345.

Lanjouw, J. O. and M. Schankerman (2001) "Characteristics of Patent Litigation: A Window on Competition," RAND Journal of Economics. **32** (1): 129-51.

Lerner, J. (1994). "The Importance of Patent Scope: An Empirical Analysis." RAND Journal of Economics **25**(2): 319-333.

Maddala, G. (1983). Limited Dependent and Qualitative Variables in Econometrics, New York, Cambridge University Press.

Matutes, C., P. Regibeau and K. Rocket (1996). "Optimal Patent Design and the Diffusion of Innovations." RAND Journal of Economics **27**(1): 60-83.

Megna, P. and M. Klock (1993). "The Impact of Intangible Capital on Tobin's 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., K. Hamilton and D. Olivastro (1997). "The Increasing Linkage Between U.S. Technology and Public Science." Research Policy **26**(3): 317-330.

Narin, F., E. Noma and R. Perry (1987). "Patents as Indicators of Corporate Technological Strength." Research Policy **16**: 143-155.

Nordhaus, W. D. (1967). The Optimal Life of a Patent. New Haven.

Pakes, A. (1986). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." Econometrica **54**(4): 755-784.

Pitkethly, R. (1999). The Valuation of Patents :A Review of Patent Valuation Methods with Consideration of Option Based Methods and the Potential for Further Research, Judge Institute Working Paper WP 21/97. Cambridge/UK.

Priest, G. and B. Klein (1984). "The Selection of Disputes for Litigation." Journal of Legal Studies **8**: 1-56.

Reitzig, M. (2002a). Improving Patent Valuation Methods for Management – Validating New Indicators by Understanding Patenting Strategies, LEFIC WP 2002-9.

Reitzig, M. (2002b). Die Bewertung von Patentrechten – eine theoretische and empirische Analyse aus Unternehmenssicht, Deutscher Universitäts Verlag, Wiesbaden.

Reitzig, M. (2002c). Valuing Patents and Patent Portfolios from a Corporate Perspective – Theoretical Considerations, Applied Needs, and Future Challenges”, expert report OPA/CONF.1/2002/4 for the United Nations Economic Commission for Europe (UNECE).

Reitzig, M. (2003). What Determines Patent Value – Insights from the Semiconductor Industry, Research Policy**,** **32** (1): 13-26.

Rosenkopf, L. and Nerkar, A. (2001). Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry, Strategic Management Journal, 22: 287-306.

Scherer, F. M. (1965). "Firm Size, Market Structure, Opportunity and the Output of Patented Inventions." American Economic Review **December**(55): 1097-1125.

Schmoch, U. and N. Kirsch (1994). Analysis of International Patent Flows, Karlsruhe, Fraunhofer Institut für Innovationsforschung and Systemtechnik Karlsruhe.

Scotchmer, S. and J. Green (1990). "Novelty and Disclosure in Patent Law." RAND Journal of Economics **21**(1): 131-146.

Teece, D. (1986). "Profiting from Technological Innovation: Implications for Integrating, Collaboration, Licencing and Public Policy." Research Policy **15**: 285-305.

Tong, X. and J. D. Frame (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.

Trajtenberg, M., Henderson, R. and Jaffe, A.. (1997). University versus corporate patents: A window on the basicness of invention, Economics of Innovation and New Technology, **5**: 19-50.

Verspagen, B. and I.D. Loo. (1999). Technology Spillovers between Sectors and over Time, Technological Forecasting and Social Change, **60**(3): 215-235.

Waldfogel, J. (1998). "Reconciling Asymmetric Information and Diverging Expectations Theories of Litigation." Journal of Law and Economics **49** (2): 451-476.

Appendix A

Validation Studies of Patent Value Indicators in the Reduced Form

Indicator	Study (Authors)	Methodology			Results (Significance/Sign)
		- **Number of Observations / Industrial Sector**	- **Dependent Variable (if Multivariate Test)**	- **Type of Validity**	
Backward Citations[#]	Carpenter, Cooper et al. (1980)	- 399 US patents on prostaglandines	- None	- Face validity	-
	Narin et al. (1987)	- 17 Pharmaceutical companies	- Indicators of corporate technological strength	- Construct validity	- Patent citations to the scientific literature significant - Positive correlation
	Lanjouw and Schankerman (2001)	- 10,378 US patents	- Infringement and challenge suits	- Construct validity	- Significant - Negative correlation
	Lanjouw and Schankerman (1999)[♣]	- Approx. 8000 US patents	- Latent variable of patent quality	- Construct validity	- Forward citations strongest predictor of patent quality, claims and backward citations second, family size forth (all coefficients positive)
	Harhoff et al. (1999)	- 57,782 observations on 778 DE patents	- Patent value	- Direct validity	- Significant - Positive correlation
	Harhoff and Reitzig (2002)	- 13,389 EP patents	- Opposition	- Construct validity	- Insignificant (when not including international search references) - Significant (when including international search references)
	Reitzig (2002a)	- 813 EP patents	- Opposition	- Construct validity	- Patent references significant - Positive
Forward Citations	Narin et al. (1987)	- 17 Pharmaceutical companies	- Sum of the means of six financial variables	- Construct validity	- Significant - Positive
	Trajtenberg (1990)	- N>2,000 (sales of CT scanners to US hospitals)	- Social value computed for the innovation	- Construct validity	- Significant - Non-linear (convex)
	Lanjouw and Schankerman (1999)[♣]	- Approx. 8000 US patents	- Latent variable of patent quality	- Construct validity	- Forward citations strongest predictor of patent quality, claims and backward citations second, family size forth (all coefficients positive)
	Albert et al. (1991)	- 77 Patents	- Patent value	- Direct validity	- Significant - Positive - Non-linear
	Harhoff et al. (1999)	- 57,782 DE patents	- Patent value	- Direct validity	- Significant

	Harhoff and Reitzig (2002)	- 13,389 EPO patents	- Opposition	- Construct validity	- Significant - Non-linear (convex)
	Hall et al. (2000)	- 17,111 US Manufacturing Patents	- Tobin's Q	- Construct validity	- Significant - Positive correlation
Family Size	Lanjouw et al. (1998)	- N>20,000 DE patent renewals	- Renewal decision	- Construct validity	- Significant
	Lanjouw and Schankerman (1999)♣	- approx. 8000 US patents	- Latent variable of patent quality	- Construct validity	- Forward citations strongest predictor of patent quality, claims and backward citations second, family size forth (all coefficients positive)
	Guellec and van Pottelsberghe de la Potterie (2000)	- 23,487 EP patent applications	- Patent grant	- Construct validity	- Significant - Positive for G3 patenting
	Harhoff and Reitzig (2002)	- 13,389 EP patents	- Opposition	- Construct validity	- Significant - Positive correlation
	Reitzig (2002a)	- 813 EP patents	- Opposition	- Construct validity	- Significant - Positive
Scope	Lerner (1994)	- 535 Venture financed bio-technology firms	- Value of the firm	- Construct validity	- Significant - Linear / positive correlation
	Harhoff et al. (1999)	- 57,782 DE patents	- Patent value	- Direct validity	- Insignificant
	Harhoff and Reitzig (2002)	- 13,389 EP patents	- Opposition	- Construct validity	- Insignificant
	Lanjouw and Schankerman (2001)	- 10,378 US patents	- Infringement and challenge suits	- Construct validity	- Significant for infringement suits - Negative correlation
	Reitzig (2002a)	- 813 EP patents	- Opposition	- Construct validity	- Insignificant
Ownership	Lanjouw and Schankerman (2001)	- 10,378 US patents	- Infringement and challenge suits	- Construct validity	- Significant - Negative effect of individual ownership
	Harhoff and Reitzig (2002)	- 13,389 EP patents	- Opposition	- Construct validity	- Significant - Negative effect of individual ownership
	Guellec and van Pottelsberghe de la Potterie (2000)	- 23,487 EP patent applications	- Patent grant	- Construct validity	- Cross-border ownership significant - Positive
Legal Argument	Harhoff et al. (1999)	- 57,782 observations on 778 DE patents	- Patent value	- Direct validity	- Significant - Positive
Patenting Strategy	Guellec and van Pottelsberghe de la Potterie (2000)	- 23,487 EP patent applications	- Patent grant	- Construct validity	- Significant - PCT II strongly positive
	Reitzig (2002a)	- 813 EP patents	- Opposition	- Construct validity	- PCT II significant - Positive

Number of Applicants	Guellec and van Pottelsberghe de la Potterie (2000)	- 23,487 EP patent applications	- Patent grant	- Construct validity	- Significant - Negative
Number of Cross-Border Research Co-operations	Guellec and van Pottelsberghe de la Potterie (2000)	- 23,487 EP patent applications	- Patent grant	- Construct validity	- Significant - Positive
Accelerated Examination Request	Reitzig (2002a)	- 813 EP patents	- Opposition	- Construct validity	- Significant - Positive
Claims	Tong and Frame (1992)	- 7,531 US patents	- R&D - GNP	- Construct validity	- Significant - Claim counts outperform patent counts
	Lanjouw and Schankerman (1999)♣	- approx. 8000 US patents	- Latent Variable of Patent Quality	- Construct validity	- Forward citations strongest predictor of patent quality, claims and backward citations second, family size forth (all coefficients positive)
	Lanjouw and Schankerman (2001)	- 10,378 US patents	- Infringement and challenge suits	- Construct validity	- Significant - Positive
	Reitzig (2002a)	- 813 EP patents	- Opposition	- Construct validity	- Dependent and independent product claims significant and positive - Application claims significant and negative

Note: #: Note that several indicators compiled from backward citations are summarized in this table, e.g. science linkage, patent references or legal quality.
This table contains a selection of studies and is not complete.
♣: In the first part of their paper, Lanjouw and Schankerman (1999) estimate a latent variable construct for patent quality and assume that forward citations, backward citations, family size and claims contribute to patent value; the results are therefore somehow ‘self-referential’. The second part (verification of the results using renewal and litigation data) is not discussed here.

Appendix B: Derivation of the ML estimator

Condition (1) for an opposition being filed is as follows:

$$p_{rj} \cdot p_{opponent}(rejection|opp.) + p_{am} \cdot p_{opponent}(amendment|opp.) + p_{rv} \cdot p_{opponent}(revocation|opp.) > p_{no} \quad (1)$$ [24]

Using assumptions A1, A2 and A4 and abbreviating $p_{Opponent}$ to p, this yields equation (3a)

$$p \cdot \{-c - g(novelty, inventive\ activity, other\ factors)\} > (0 + c) \quad (3a)$$

that denotes the simplified condition for an opposition being filed.
Defining

$$novelty, inventive\ step = xb + e \quad (3b)$$

and accordingly

$$other\ factors = zg + h \quad (3c)$$

yields

$$g(novelty, inventive\ step, other\ factors) = xb + zg + e + h \quad (3d)$$

as the simplest form for g (linear combination).

If H1, H2, and H3 are true, condition (3) must then take the following form:

$$p[-c - (xb + e + zg + h - c)] > (0 - (xb + e + zg + h - c))$$
$$\Leftrightarrow (1 - p)(xb + e + zg + h) - c > 0$$

$$\Leftrightarrow (1 - p)(xb + zg) - c > (1 - p)(e + h) \quad (5).$$

In the following, the observable term $(1 - p)(xb + zg) - c$ is referred to as a, and the unobservable term $(1 - p)(e + h)$ is referred to as x.

A likelihood function reflecting Equation (5) has three distinct probabilities, *p(W=0)*, *p(W=1,V=0)*, and *p(W=1,V=1)*, where

---

24 Numbering consistent with formula numbering in the text.

W=1 denotes the occurrence of an opposition (W=0 otherwise), and
V=0 denotes the revocation of the patent (V=1 otherwise).

According to (5) the likelihood of no opposition taking place is given by

$$p(W = 0) = \Phi\left(-\frac{a}{s_x} \right) \qquad (6),$$

assuming that e and h are normally distributed with covariance $\begin{bmatrix} 1 & r_{e,h} \\ r_{e,h} & 1 \end{bmatrix}$.

In the case of an opposition there are two possible outcomes depending on the (opponent's anticipation of the) EPO's decision. Using assumption A3 yields the following additional condition for the opponent's anticipated probability of the patent being upheld after the opposition (2^{nd} stage):

$$p = P(xb + e) > 0 \qquad (7).$$[25]

The probability of an opposition taking place and the patent being upheld (or amended) is described by the binormal distribution

$$p(W = 1, V = 1) = \Phi^2\left(\frac{a}{s_x}, xb, r_{x,e} \right) \qquad (8),$$

where $s_x^2 = 2(1 + r_{h,e})\ if\ s_h = 1,\ and\ r_{x,e}$ is the correlation coefficient between the disturbances of the 1^{st} stage (opposition yes/no) and 2^{nd} stage (patent revoked/upheld or amended) that can be calculated to

$$r_{x,e} = \frac{(1 - \Phi(xb)) \cdot (1 + r_{h,e})}{(1 - \Phi(xb)) \cdot \sqrt{(2 + 2 \cdot r_{h,e})}} = \sqrt{\frac{1 + r_{h,e}}{2}} \qquad (9),$$

25 Note: this estimator does not treat the anticipated probability of an EPO ruling as a *conditional* probability given that opposition actually takes place. This simplification is, however, vindicated easily considering that *potential* opponents would *always* have an anticipated probability of winning or losing their case – *no matter* whether they eventually file an opposition or not. In that respect it is consistent with one of the major underlying assumptions of the paper (A3). Statistically speaking, the simplification has advantages in that it keeps the complexity of the estimator manageable.

Analogously to (9), the probability of the patent being revoked can be calculated as

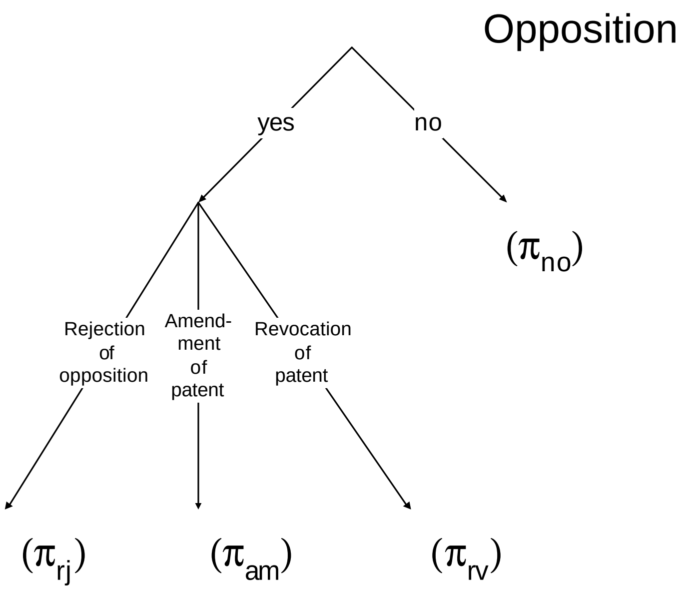
$$p(W=1, V=0) = \Phi^2\left(\frac{a}{s_x}, -\sum xb, -r_{x,e}\right) \qquad (10).$$

Equations (6), (8), and (10) finally constitute the ML estimator function (4):

$$\log L = (1-W)\cdot\log\left(\Phi\left(-\frac{a}{s_x}\right)\right) + W\cdot\left[V\cdot\log\left(\Phi^2\left(\frac{a}{s_x}, \sum xb, r_{x,e}\right)\right) + (1-V)\cdot\log\left(\Phi^2\left(\frac{a}{s_x}, -\sum xb, -r_{x,e}\right)\right)\right].$$

Figure 1

Decision Tree of the Opponent



<u>Legend:</u>

π_{no}: Profits of the 'opponent' in the case of no opposition (=0)

π_{rj}: Profits of the opponent if the opposition is rejected (=-c)

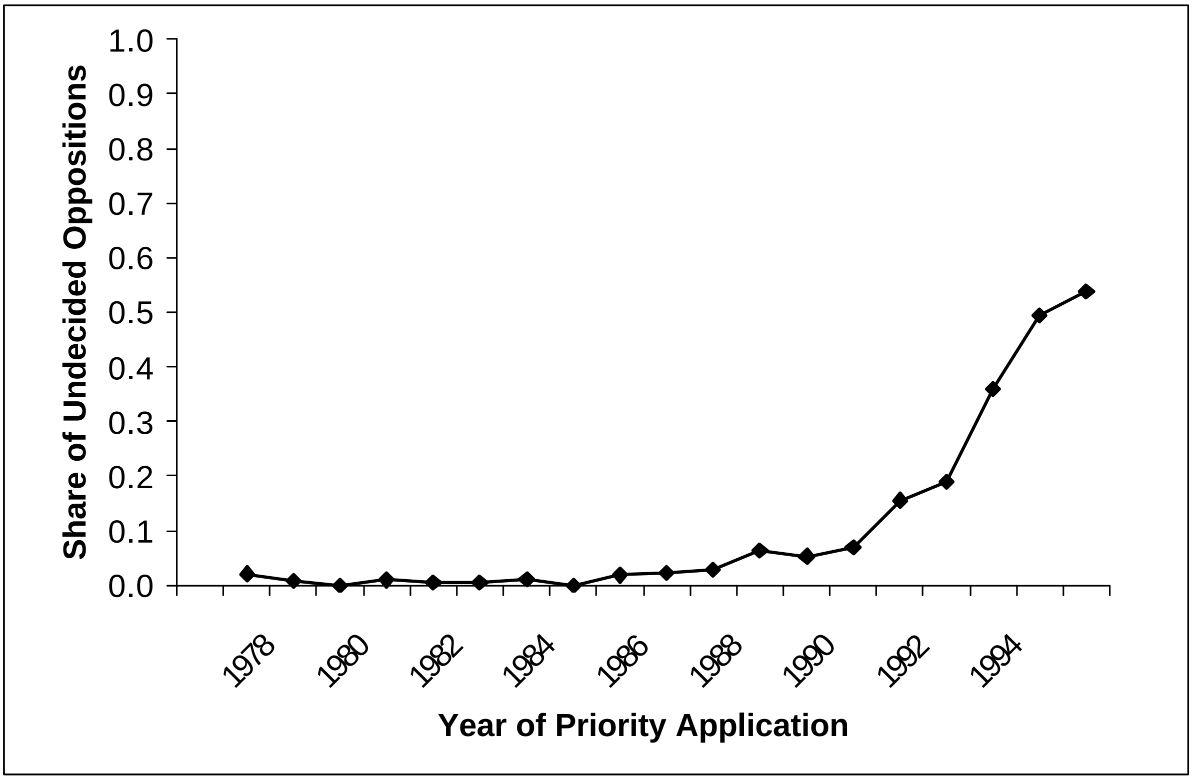
π_{am}: Profits of the opponent if the patent is amended (=-c)

π_{rv}: Profits of the opponent if the patent is revoked

(=f(novelty, inventive step, other factors, c))

Figure 2

Share of Unidentified Opposition Outcomes[26] *Among All Opposition Cases in Polymers vs. Year of Priority Application*



26 As can be reconstructed from electronic sources. See the text for details.

Table 1

Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Opposition (1: yes, 0: no)[1)]	0.12		0	1
Rejection of Opposition (1: yes, 0: no) [2)]	0.26		0	1
Amendment after Opposition (1: yes, 0: no) [2)]	0.28		0	1
Revocation of Patent after Opposition (1: yes, 0: no) [2)]	0.39		0	1
Opposition Procedure Closed (1: yes, 0: no) [2)]	0.05		0	1
Opposition Outcome not Definable (1: yes, 0: no) [2)]	0.02		0	1
Number of Backward Citations to the Patent Literature (incl. international search) [1)]	3.48	1.90	0	24
Number of Backward Citations to the Non-Patent Literature (incl. international search) [1)]	0.60	0.93	0	13
Number of Designated States (Family Size) [1)]	7.46	3.20	1	15
Number of Applicants[1)]	1.04	0.23	1	5
Number of Inventors[1)]	2.72	1.47	1	18
Number of Forward Citations (3-year frame) [1)]	0.66	1.07	0	18
Accelerated Examination Request (1: yes, 0: no) [1)]	0.01		0	1
PCT (1: yes, 0: no) [1)]	0.08		0	1
Time dummy (priority dates1978-1980)	0.15		0	1
Time dummy (priority dates1981-1985)	0.35		0	1
Time dummy (priority dates1986-1990)	0.50		0	1

Legend: 1): Entire sample comprising N=17,123 patents.
2): Sample of opposed patents comprising N=2,034 patents. Outcomes computed electronically as described in FN 12.

Table 2

Probability Estimations in the Reduced Form

A. Likelihood of an Opposition

B. Likelihood of Patent Maintenance following Opposition

Variable[3]	**Column A (S.D.)**[1]		**Column B (S.D.)** [2]	
	2.A.I.	**2.A.II.**	**2.B.I.**	**2.B.II.**
Number of Backward Citations to the Patent Literature	0.28*** (0.06)	0.24*** (0.06)	0.11 (0.14)	0.13 (0.14)
Number of Backward Citations to the Non-Patent Literature	-0.89*** (0.14)	-0.58*** (0.14)	0.26 (0.35)	0.17 (0.36)
Number of Designated States (Family Size)	0.14*** (0.04)	0.32*** (0.04)	0.15* (0.09)	0.07 (0.09)
Number of Applicants	-0.14 (0.56)	0.18 (0.57)	-0.60 (1.31)	-0.72 (1.31)
Number of Inventors	-0.02 (0.08)	0.17** (0.08)	0.48*** (0.20)	0.41** (0.20)
Number of Forward Citations (3-year frame)	0.51*** (0.09)	0.71*** (0.09)	0.44** (0.22)	0.36* (0.22)
Accelerated Examination Request (1: yes, 0: no)	1.19 (1.36)	2.84** (1.40)	3.11 (3.12)	2.42 (3.12)
PCT (1: yes, 0: no)	0.13 (0.54)	0.67 (0.55)	-1.60 (1.24)	-1.98 (1.26)
Time dummy (priority dates 1981-1985)	-[4]	-2.76 (0.35)	-	0.43 (0.71)
Time dummy (priority dates 1986-1990)	-	-6.76*** (0.36)	-	2.11*** (0.80)
Constant	-13.53*** (0.73)	-11.97*** (0.77)	0.15 (1.69)	0.29 (1.74)
LR c^2 *–test (8;10;8;10)*	*101.65*	*487.62*	*17.33*	*25.63*
P-Value	<0.01	<0.001	=0.03	<0.01

Legend:
1): Entire sample comprising N=17,123 patents.
2): Sample of opposed patents comprising N=2,034 patents.
3): All coefficients are multiplied by factor 10.
4): Coefficient not tested in the respective specification.

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

Table 3

Probability Estimations in the Structural Form
XB: Indicators of Technical Value Drivers
ZG: Indicators of Non-technical Value Drivers

Variable		**A Column (S.D.)**	**B Column (S.D.)**
XB[1)]	Number of Backward Citations to the Non-Patent Literature	-0.37** (0.15)	-0.55*** (0.16)
	Number of Applicants	-0.04 (0.23)	0.04 (0.59)
	Number of Inventors	-0.00 (0.04)	0.23** (0.11)
	Number of Forward Citations (3-year frame)	0.64*** (0.17)	0.73*** (0.13)
	Time dummy (priority dates 1981-1985)	-[2)]	-0.20 (0.70)
	Time dummy (priority dates 1986-1990)	-	0.34 (0.78)
	Constant	-2.42 (1.62)	-3.66*** (0.98)
ZG[1)]	Number of Backward Citations to the Patent Literature	0.10*** (0.04)	0.16*** (0.04)
	Number of Designated States (Family Size)	0.05** (0.02)	0.23*** (0.05)
	Number of Forward Citations (3-year frame)	-0.33* (0.21)	0.12 (0.19)
	Accelerated Examination Request (1: yes, 0: no)	0.36 (0.48)	1.91* (1.03)
	PCT (1: yes, 0: no)	0.07 (0.19)	0.50 (0.38)
	Time dummy (priority dates 1981-1985)	-	-1.87*** (0.59)
	Time dummy (priority dates 1986-1990)	-	-5.16*** (0.87)
ρ		0.93*** (0.04)	-0.75*** (0.07)
Costs[1)]		1.36 (0.92)	3.18*** (0.84)
s_h		-0.96*** (0.30)	1.02*** (0.05)
Wald c^2 *(9;13)*		*21.28*	*102.41*
P-Value		=0.01	<0.001

Legend: Entire sample comprising N=17,123 patents.
1): Coefficient multiplied by factor 10.
2): Coefficient not tested in the respective specification.
*/**/*** Significant at 10%/5%/1% level (two-tailed tests)