Reducing uncertainty in the patent application procedure – insights from invalidating prior art in European patent applications

Christian Sternitzke[*,1,2]

[1]Ilmenau University of Technology, PATON – Landespatentzentrum Thüringen, PF 100 565, 98684 Ilmenau, Germany

[2]University of Bremen, Institute for Project Management and Innovation (IPMI), Wilhelm-Herbst-Strasse 12, 28359 Bremen, Germany

Abstract:

Achieving patent protection for an invention is a costly procedure. The patenting process in front of the patent office itself is frequently associated with substantial uncertainty about the outcome. This paper aims to identify measures to reduce this uncertainty and increase efficiency in patenting by investigating so-called world patent applications in chemicals, pharmaceuticals, and biotechnology that were subsequently not granted by the European Patent Office (EPO). Examination reports of these documents were searched for references that were regarded to invalidate novelty or inventiveness of the patent application. The results from studying these references provide several implications for patent applicants in chemicals, pharmaceuticals, and biotechnology: novelty is in many cases anticipated in patent literature. Dispersed nonpatent literature frequently anticipates the inventive step. Patent searches in the same 4-digit IPC class as the original

[*] Corresponding author: christian.sternitzke@tu-ilmenau.de, phone: +49 3677 69 4537, fax: +49 3677 69 4538

invention reveal the majority of all relevant prior art in patents. Furthermore, inventors and applicants were aware of a considerable share of invalidating prior art. This can be explained by a gambling behavior in the patenting procedure of large firms, while particularly SMEs encounter difficulties in evaluating both novelty and inventive step of their "invention".

Key words: patent and nonpatent references, citations, novelty, inventive step, non-obviousness, patent search

This is the accepted version of the paper. The final version is available under https://doi.org/10.1016/j.wpi.2008.04.007

1. Introduction

Patent applications are filed to appropriate the returns of successful and costly research and development (R&D). The outcome of the patenting procedure is a granted patent providing the applicant with the right to prevent others from using the invention. The overall patent lifecycle is associated with substantial risk and uncertainty, as illustrated in Figure 1.

{insert Figure 1 about here}

Over the whole lifecycle there are several business risks and uncertainties associated with the patented invention. On the one hand the patent may not be commercially valuable (1). A prominent example for such a situation is the pharmaceutical industry where patents involve new chemical substances that, during the clinical stage of product development, may turn out to show severe side effects on the human body and hence, are not commercially valuable. Imitability and substitutability face every patented technology and are discussed at length in the management literature. In addition to these business risks and uncertainties, there are various legal ones. Before the patent is granted (2), the applicant has to maneuver through several pitfalls in front of the patent office, thereafter (3) he faces the risk of opposition (in some countries) and the risk of patent litigation. While the latter issue is addressed frequently in the literature (see e.g. [1-3]), the former receives little attention from a patent information perspective.

During the patenting procedure, patent examiners analyze patent applications according to valid patent laws. This means that the invention needs to fulfill certain criteria such as novelty, the involvement of an

inventive step (i.e. the invention may not have been obvious for "the skilled man in the art" or in other words – trivial), the invention must be commercially applicable. Several areas are excluded from patentability, such as algorithms, the treatment of the human or animal body, etc. Additionally, the application needs to comply with several formal standards, such as the format of the submitted application documents, fees have to have been paid, and the invention has to be fully disclosed so that the "skilled man in the art" can replicate what shall be protected. The latter point is one fundamental aspect of the patent system – disclosure to stimulate innovation for legal protection that guarantees, from the legal perspective, an appropriate return on investment of the underlying R&D effort. Hence, to obtain a granted patent, applicants have to comply with all these aspects. On the one side, the applicant has to prove novelty and inventive step by describing the background of the invention where frequently references are made to prior art. These references should not contradict novelty and inventive step. Patent examiners, on the other side, scan the relevant literature, mainly patents, in order to find prior art documents that anticipate the invention or give indication regarding the inventive step. If the examiner reaches the conclusion that parts of the claimed invention do not fulfill the criteria of patentability, she discusses the patent applications' claims and cites relevant sources.

This paper aims to elicit how applicants can reduce risk and uncertainty associated with prior art of their patents. The central question is: *how can applicants be assured of meeting novelty and inventive step?* In this context examination reports of patent applications which were not granted at the

European Patent Office (EPO) are investigated with respect to references made by the examiner to prior art that anticipated either novelty or the inventive step of the invention, hence preventing the examiner from granting the patent. In this context, we try to answer the following questions:

- How important are patent versus nonpatent references?
- Which matter more for novelty and inventive step?
- Can applicants know invalidating prior art, and *do* they?
- Where to search effectively for relevant patents?

This paper is organized as follows: in the next section we briefly describe the dataset and methodology. Section three discusses the results, conclusions follow in section 4.

2. Methodology and dataset

As pointed out in the previous section, both novelty and inventive step can be assessed through analyzing the references contained in patent documents, a field known as (patent) citation analysis. Many papers have been published on patent references, following the pioneering work of Narin and his colleagues (e.g. [4-9]). In general, the literature distinguishes between patent references contained in the full-text of the patent application, so-called inventor citations originating from the inventor or the patent applicant, and so-called examiner citations coming from the patent searcher or examiner at the patent offices. These types of references are generally

available to the public, either within the published patent applications, in search reports, or in the reference section of granted U.S. patents.

It is well known that these references are quite noisy due to different policies at the patent offices: in the United States, patent law requires applicants to submit a list with prior art that might be relevant. Frequently, examiners include these into the reference list. The European Patent Office (EPO) follows the policy to cite only a minimum number of references (for a discussion, see [10-13]). Thus, only a subset of all references in patents is in fact relevant to assess novelty and the inventive step. These relevant references are not necessarily contained in search reports of European patents, but in the corresponding examination reports which, in many cases, are more exhaustive. Analyses of patent citations that explicitly take into account the references from examination reports have not been reported in the literature so far. For this paper, therefore, examination reports of the EPO are studied to elicit which references exactly invalidated the patent application by anticipating novelty or inventiveness. Invalidating in this case means that the references were relevant prior art for some claims in the patent, not necessarily 100 percent of the claims contained therein. The patent documents studied were all patents filed via the World Intellectual Property Organization's (WIPO) Patent Cooperation Treaty (PCT). These so-called WO patent applications aimed at a variety of countries and, hence, belong to particular important inventions. More specifically, these WO patent applications had to have the EPO as designated office, relate to the field of chemicals, pharmaceuticals, and biotechnology, and had to have failed in the examination procedure. Failure could both have led to an

official rejection of the patent application, or to withdrawal of the application through the applicant, which in practice frequently occurs when the patent office had communicated severe objections against the patent application [14]. The sample was drawn from a set of PCT patent applications investigated in an earlier publication (see [15]) that analyzed all PCT applications with priority dates from the first half of December 1996, in total 2,719 applications. Among them, 2,600 had the EPO as designated office, and after deleting sets with missing data, 2,564 applications remained. Among these 2,564 applications 171 belonged to the technology fields investigated, namely organic fine chemicals; macromolecular chemistry, polymers; pharmaceuticals, cosmetics; and biotechnology, as defined via classes of the International Patent Classification (IPC) by [16]. These 171 patent applications were searched in EPOLINE, the European patent register. 92 patent applications had to be excluded, either because no electronic file was available, the electronic file was incomplete, or no examination had taken place. A handful of patent examinations had to be excluded because either no proper examination could have been carried out due to obscurities relating to the priority date, or the prior art mentioned related to earlier, novelty destroying disclosures of the same applicant. The remaining 79 patent applications formed the basis for this paper. In many of them, invalidating prior art was not the only reason for failure.

For further analyses, patent family and classification information was retrieved from the Derwent World Patents Index (WPINDEX) database. Where no information on the patent family's main class was available in WPINDEX, the first IPC class mentioned in the Esp@cenet database was

chosen. In the case where inventor's references were analyzed, the full text of the WO patent applications was screened manually. Finally, when the WO document was in Japanese, the corresponding European patent application was scanned.

3. Results and Discussion

Of the 79 WO patent applications identified, 20 had failed in the examination procedure even though no invalidating prior art was mentioned, neither regarding novelty nor the inventive step. Here, obviously other objections by the patent office led to rejection or withdrawal. The remaining 59 WO patent applications contained both patent as well as nonpatent references from the applicant, and only a subset was, in fact, regarded to be invalidating by the examiner. Nonpatent references mainly refer to scientific articles/publications, in many cases from prestigious journals, in a few cases also to books. They did not seem to be entirely different in nature than the nonpatent literature mentioned in search reports or the references cited-section of US patents as discussed in e.g. [17-18]. Table 1 gives an overview about the 59 patent documents and the references contained therein.

{insert Table 1 about here}

3.1 Importance of patent versus nonpatent references

How important are patent versus nonpatent references? Table 1 reveals that 80 percent of all patent applications under examination contained at least one invalidating patent reference, thus patent literature, as one would expect, plays a crucial role when searching for prior art. However,

nonpatent literature may not be ignored: 59 percent of the patent applications contain at least one reference to nonpatent literature, and 20 percent contain nonpatent references exclusively. Hence, a priori searches by applicants in chemicals, pharmaceuticals, and biotechnology need to focus on both.

3.2 Patent and nonpatent references vs. novelty and inventive step

From a patent searchers' perspective it is interesting to know if patent or nonpatent literature matter more for novelty or the inventive step. Therefore, in a next step, it was investigated why these references had been incorporated. The basis for this analysis was a manual screening and assessment of the examination reports to distinguish between references relating to novelty (coded as N) or the lack of the inventive step (coded as I)? In this context it was further assessed if the reference was invalidating itself, or only together with other references. This characterization can also be found in EPO or WIPO search reports: X references relate to a document that is of particular relevance itself, while Y coding refers to documents that are particularly relevant when taken together with other references. It was found that about half of all WO patent applications in the sample contained both N and I coded references and a quarter each only N and only I. Table 2 and 3 provide a more detailed overview about the results.

{insert Table 2 and 3 about here}

In total, 125 patent references were contained within the 59 examination reports, and 86 nonpatent references. About two thirds of all patent references relate to NX documents, signifying that these documents

contained prior art that invalidated novelty of at least one claim of the patent application. About a quarter of the patent references refer to IX and IY documents each. In contrast, nonpatent references contain fewer documents relevant for rendering novelty of claims obsolete; they seem to be more important for assessing the inventive step. Checking for significance by means of a chi-square test, it was found that patents are significantly more utilized to assess novelty than nonpatent publications, but the latter are used significantly more than expected to evaluate the existence of an inventive step. Hence, these findings seem to reflect search strategies at the patent office: if examiners do not find the complete disclosure of a patent claim as novelty-related information in a patent document, they tend to search for single pieces of information across various sources, in particular, nonpatent literature.

3.3 Availability of invalidating prior art to applicants

With the findings of the previous section in mind: can applicants know invalidating prior art, and *do* they? Figure 2 provides an overview about the age structure of the references to answer the first part of the question. In each case, the publication year of the reference was counted.

{insert Figure 2 about here}

The results show that more than 80 percent of all references were theoretically available to the applicant since they were published prior to the priority date. Figure 2 also reveals that 50 percent of the references originate from the last four years. Those references published after the priority date relate in the case of patent references to patent applications that had been

filed earlier at the EPO than the applications under consideration, i.e. they had an earlier priority date and thus were regarded as prior art. Obviously, the applicants could not have been aware of these documents. Surprisingly, examiners cited also articles that had been published after the priority date, with the inference in some cases that they violated novelty. The examiners not always gave reasoning for taking these into account. However, in one case it was mentioned that the publications were very descriptive in nature. So it could be that – even though these nonpatent references had not been characterized as review articles in the Science Citation Index – they might refer to various sources that are invalidating in nature for the patent application, but were not cited individually in order to minimize the total number of references in the examination report. Another phenomenon that can be observed in Figure 2 is that referenced patents in general range back further than publications, which can be explained with data availability: patent documents are accessible more easily for a longer time period than nonpatent literature.

So since many invalidating references were theoretically accessible for the applicant, the question arises if he or she (or the patent attorney formulating the patent application) was in fact aware of some invalidating prior art. It was found for front-page references in granted US patent documents that 70 percent of all references were known by the inventors at the point the patent application was submitted to the patent office [19]. This does not necessarily mean that these references are also invalidating. Otherwise it is unlikely that the inventors would have applied for the patent. Comparing references found in the full text of the patent application with references of

invalidating prior art from examination reports (while controlling for patent family documents) reveals the answer to the question posed: yes, to a certain degree applicants knew about invalidating prior art!

On the basis of references, applicants were aware of 20 percent of all patent references, and 31 percent of all nonpatent references. When taking the patent applications as a basis where the references are contained therein, Table 4 demonstrates that the applicant was aware of at least one invalidating patent reference in 30 percent of all WO patent applications, and he or she knew about at least one invalidating nonpatent reference in almost 50 percent of all cases. In 15 percent of all WO patent applications the applicant was even aware of all invalidating references. Could this structure be rooted in different risk attitudes of applicants? Small and medium-sized companies (SMEs) possess, in general, less financial resources than large firms. Hence, it could be that large firms pursue a more aggressive patenting strategy than smaller firms, implying that large firms maybe know about prior art but rather take the risk to gamble?[†] In order to answer this question, we related firm size to knowledge about prior art, taking the definition for SMEs of the European Commission as a basis.[‡] Table 5 presents the results. Large firms know about significantly more invalidating patent references as one would expect, while SMEs are aware of a high share of invalidating nonpatent references. Since large firms tend to utilize patent information to a much higher extent than SMEs [20], one

[†] The author is grateful to Adam Bartkowski for pointing towards this relationship.

[‡] Independent inventors as well as universities were also counted as SMEs since they, in most cases, possess less financial resources than large firms. In one case, the US Navy was an applicant and, in this case, treated as large firm.

can assume that large firms showed the tendency to gamble. SMEs, in contrast, seem to encounter difficulties in assessing to what degree prior art is, in fact, invalidating. In other words, they have a lower understanding to what degree they may combine knowledge from various sources without running into trouble with the nonobviousness criterion at the patent office.

{insert Tables 4 and 5 about here}

Is there any difference between recognizing novelty or inventive step from the applicant's perspective? Table 6 answers this question: known nonpatent references have a distribution that is, according to a chi-square test, not statistically significant from the distribution of unknown nonpatent references. For patent references, the situation is different: applicants know significantly more about references that are invalidating regarding the inventive step, and to a lower extend about novelty-related patent references. This finding clearly underlines the fuzzy definition of inventiveness, but it can not only mean that assessing novelty from patent sources is much easier from the applicant's perspective than evaluating nonobviousness, it can also imply that applicants search patent sources to a lower degree or less effectively, which would go hand in hand with the finding that applicants know less about patent than nonpatent references.

{insert Table 6 about here}

It can be concluded that the majority of invalidating prior art is not only theoretically available to the applicant, he or she also knows about a considerable share of it. This share is even higher when taking both aspects together, i.e. calculating it on the basis of the patent and nonpatent literature

that has been published prior to the patent application's priority date. Large firms tend to rather gamble on the grant of the patent while SMEs encounter difficulties in assessing nonpatent literature.

3.4 Effective search strategies for identifying invalidating prior art

How can the applicant ex ante improve the search strategy for patent literature? Since the IPC plays an important role in patent searches, the IPC main class of the WO application was compared with the IPC main class of the referenced invalidating patents. To obtain a more detailed picture, the IPC was broken down in subclass (4-digits), group (7-digits) and main- or subgroup (full number). Additionally, it was controlled for reference types such as novelty and inventive step. Results can be found in Table 7.

{insert Table 7 about here}

About 70 percent of all invalidating patent references originate from the same 4-digit IPC class as the WO application under examination, 50 percent from the 7-digit, and less than a third is assigned to the same full IPC class as the WO application. Thus a major share of invalidating prior art can be found just by searches in classes were the applicant expects his application to be assigned to. Taking the reference types as a basis, 70 percent of all NX patent references come from the same 4-digit IPC, about the equal amount from IX patent references, and app. 83% from all IY patent references. While about one third of all NX and IX patent references originate from exactly the same IPC class, only 17.2% of all IY patent references do. However, for the overall table, this finding proves not to be statistically significant in a chi-square test, but it indicates a well-known phenomenon:

combining knowledge from the core area of the invention involves a lower inventive step than combining knowledge from areas further apart. If novelty-related knowledge is also comprised in single documents, then it is very likely that they, in fact, belong to the same 7-digit IPC class, while, in the case of inventive knowledge, single pieces of relevant inventive knowledge may be distributed to a higher extent across several adjacent subclasses. To conclude, searching documents in the same IPC classes where the patent is likely to be assigned to should yield most of the relevant prior art.

3.6 Limitations

The analyses in this paper encounter several limitations. First of all, to demonstrate that some claims are not novel or inventive, European patent examiners need to show only as many references as are necessary to prove that the claims are not patentable [21]. Thus, only a subset of references cited in search reports may be mentioned as invalidating in the examination report, even though more documents from the search report might be invalidating as well. Second, some patent examiners tended to not assess the inventive step when novelty of a claim was already rejected (e.g. see examination report of WO 9825944), which is feasible because it is sufficient to prevent a patent from being granted if only one of the three patentability criteria is not met. Other patent examiners did not follow this policy (e.g. see examination report of WO 9825961) and checked both novelty and inventive step. This might have led to an overestimation of the impact of novelty within prior art documents. Third, invalidating prior art

should also be available in examination reports to patent applications that were subsequently granted, but in this case they should only relate to a subset of claims contained in the patent application. Also including findings from these sources could broaden the statistical basis of the conclusions drawn in this paper.

4. Conclusions and Implications

Developing technological innovations is a costly process. Protecting these innovations through patents is also connected with substantial costs, such as the application fee at the patent office, drafting the text of the application, maybe with support from a patent attorney, and, if the innovation appears to be particular valuable and is aimed to be registered in a number of countries, substantial translation fees of the document. Only a granted patent provides the owner with the right to exclude others from using the invention. The outcome of the patent application process, however, is associated with a substantial amount of uncertainty. This paper addressed this issue and investigated WO patent applications in the field of biotechnology, chemistry, and pharmaceuticals aiming at a variety of countries worldwide that had failed at the European Patent Office. References to relevant prior art made by the examiners were analyzed. The results have several implications for patent applicants in these fields and can help increasing efficiency in patenting: to reduce uncertainty in the patenting procedure, patent applicants should ex ante search both *recent* patent and nonpatent literature. Nonpatent references alone are accountable for about one third of all relevant references. It was revealed that patents play a crucial role to

determine novelty. In many cases the inventive step underlying the WO patent application could be anticipated through combining several nonpatent sources. Another finding is that applicants or inventors were aware of invalidating prior art in a major share of all WO patent applications. Patent documents coming from the same 4-digit IPC class anticipate the majority of all invalidating patent references. Large firms know about a significant portion of invalidating nonpatent references. Their financial power seems to induce a gambling behavior regarding the patenting processes' outcome. SMEs, in contrast, seem to face difficulties in assessing nonpatent literature, which, in general, relates rather to the inventive step of the invention.

Future research should expand the scope of this survey to other technology fields as well. The impact of nonpatent literature as invalidating prior art should, for instance, vary considerably here. It would also be of interest to investigate if text-mining techniques would be able to recognize if documents found by applicants are actually invalidating in nature or not. Such kind of analyses would be able to help save applicants further expenses during the granting procedure or ex ante help drafting the application accordingly. The recommendation to search relevant prior art in the same 4-digit IPC class poses a challenge from the patent searchers' perspective since the amount of relevant documents increases exponentially when moving from the same full IPC-class towards the subclass. Here, intelligent text-mining tools could prove to be valuable as well. It would, nevertheless, also be worth questioning applicants and inventors about their perception of the references they knew but that were subsequently recognized as invalidating by the examiner.

Acknowledgements

The author would like to thank his colleagues at PATON as well as the participants of the R&D Management Conference 2007 in Bremen for helpful discussions on an earlier draft of this paper.

References

[1] Lemley MA. Essay - Rational ignorance at the patent office. Northwestern University Law Review 2001;95:1495-1532.

[2] Harhoff D, Reitzig M. Determinants of opposition against EPO patent grants-the case of biotechnology and pharmaceuticals. International Journal of Industrial Organization 2004;22:443-480.

[3] Sherry EF, Teece DJ. Royalties, evolving patent rights, and the value of innovation. Research Policy 2004;33:179-191.

[4] Carpenter M, Cooper M, Narin F. Linkage between basic research literature and patents. Research Management 1980;13:30-35.

[5] Carpenter MP, Narin F, Woolf P. Citation rates to technologically important patents. World Patent Information 1981;3:160-163.

[6] Carpenter MP, Narin F. Validation Study: Patent Citations as Indicators of Science and Foreign Dependency. World Patent Information 1983;5:180-185.

[7] Narin F, Carpenter MP, Woolf P. Technological performance assessments based on patents and patent citations. IEEE Transactions on Engineering Management 1984;EM-31:172-183.

[8] Narin F, Noma E, Perry R. Patents as indicators of corporate technological strength. Research Policy 1987;16:143-155.

[9] Narin F, Olivastro D. Technology indicators based on patents and patent citations. In: van Raan AFJ (Eds.). Handbook of quantitative studies of science and technology, Amsterdam: North-Holland, 1988.

[10] Meyer M. Patent citations in a novel field of technology - What can they tell about interactions between emerging communities of science and

technology? Scientometrics 2000;48:151-178.

[11] Mogee MA. International patent data for technology analysis and planning. Final Report. Center of Innovation Studies, Lehigh University, Bethlehem, PA, 1991. Available from http://cims.ncsu.edu/reports/mogee_0291.pdf (as of 02-Feb-2005).

[12] Michel J, Bettels B. Patent citation analysis. A closer look at the basic input data from patent search reports. Scientometrics 2001;51:185-202.

[13] EPO. Utilization of patent protection in Europe. World Patent Information 1995;17:100-105.

[14] Lazaridis G, van Pottelsberghe de la Potterie B. The rigour of EPO's patentability criteria: An insight into the ''induced withdrawals''. World Patent Information 2007;29:317-326.

[15] Sternitzke C. The international preliminary examination of patent applications according to the Patent Cooperation Treaty – a proxy for patent value? Scientometrics 2009;in press.

[16] DTI/OST. DTI/OST Technology group & IPC Subclass mapping (as of June 27, 2006). Available from http://www.ipaustralia.gov.au/pdfs/ statistics/Technology%20Groups.doc.

[17] Hicks D, Breitzman T, Olivastro D, Hamilton K. The changing composition of innovative activity in the US - a portrait based on patent analysis. Research Policy 2001;30:681-703.

[18] Callaert J, Van Looy B, Verbeek A, Debackere K, Thijs B. Traces of Prior Art: An analysis of non-patent references found in patent documents. Scientometrics 2006;69:3-20.

[19] Jaffe AB, Trajtenberg M, Fogarty MS. The meaning of patent citations: report on the NBER/case-western reserve survey of patentees. NBER working paper series. Cambridge, Mass.: NBER, 2000.

[20] EPO. Utilisation of patent protection in Europe. EPO-Script No. 3. Munich: European Patent Office, 1994.

[21] EPO. Guidelines for Examination, Part B, Chapter IV, Search Procedure and strategy (as of 13 Oct 2005). Available from http://www.european-patent-office.org/legal/gui_lines/e/b_iv_3_1.htm.

Figure 1: Risk and uncertainty over the patent's lifecycle

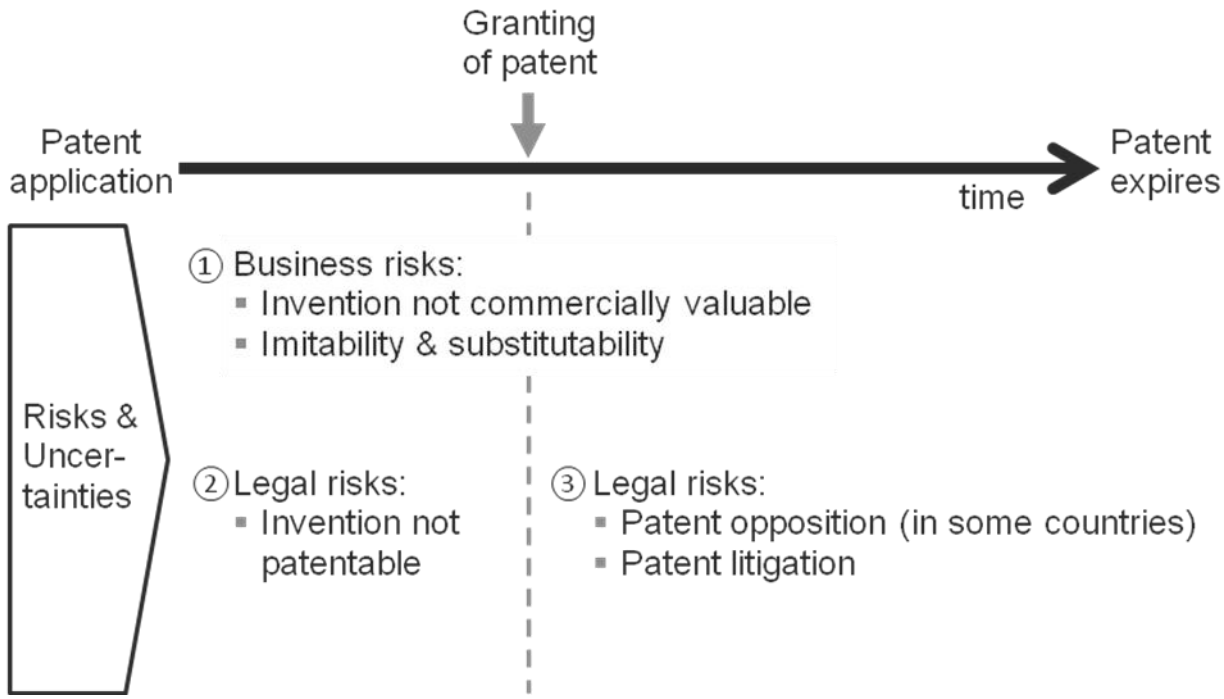

Table 1: Distribution of references contained in examination reports among patents and nonpatent references

Distribution of references		Only patent references	Only nonpatent references	Total
a) Number of patent references >=1	47 (80%)		12 (20%)	59 (100%)
b) Number of nonpatent references >=1	35 (59%)	24 (41%)		59 (100%)
a) AND b)	23 (39%)			

Table 2: Distribution of references' purpose:

Type of reference	Patent references[§]	Nonpatent references[§]
Total	125 (100%)	86 (100%)
NX	80 (64.0%)	39 (45.3%)
NY	2 (1.6%)	2 (2.3%)
IX	35 (28.0%)	26 (30.2%)
IY	29 (23.2%)	32 (37.2%)

[§] more than one entry per reference possible.

Table 3: Chi-square test of the distribution of references' purpose

Type of reference	Number of references	Patent references[§]	Nonpatent references[§]
NX	Observed	80	39
	Expected	70.9	48.1
NY	Observed	2	2
	Expected	2.4	1.6
IX	Observed	35	26
	Expected	36.4	24.6
IY	Observed	29	32
	Expected	36.4	24.6
Total		125	86

[§] more than one entry per reference possible.

p=0.077

Figure 2: Age distribution of references (τ-1 implies that the reference was published one year prior to the priority year of the patent application).

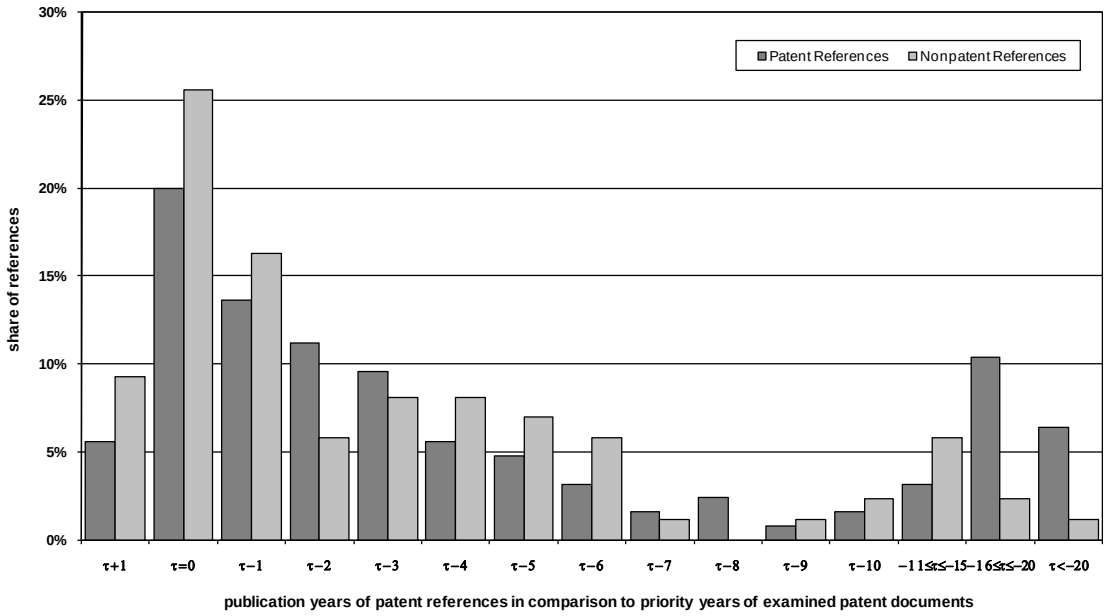

Table 4: Patent applications where the applicant was aware of references

Distribution of references	Total sample	Applicant or attorney was aware of	
		some references	all references
Number of patent references >=1	47 (100%)	14 (30%)	
Number of nonpatent references >=1	35 (100%)	16 (46%)	
Number of patents with invalidating prior art	59 (100%)		9 (15%)

Table 5: Chi-square test for reference types and firm size

Reference	Number of references	SME	Large firm
Patent reference	Observed	4	10
	Expected	7.5	6.5
Nonpatent reference	Observed	12	4
	Expected	8.5	7.5

p=0.011

Table 6: Chi-square test for the distribution of references' purpose among references known and unknown from the applicant

Type of reference	Number of references	Patent references[§]		Nonpatent references[§]	
		Unknown by applicant	Known by applicant	Unknown by applicant	Known by applicant
NX	Observed	69	11	27	12
	Expected	61.9	16.8	26.0	12.8
NY	Observed	2	0	2	0
	Expected	1.5	0.4	1.3	0.5
IX	Observed	19	16	15	11
	Expected	27.1	9.4	17.3	9.3
IY	Observed	23	6	22	10
	Expected	22.4	6.5	21.3	10.5
Total		100	25	59	27

[§] more than one entry per reference possible.

For patents: p=0.002; for nonpatent references: p=0.548

Table 7: Overlap of patent references with IPC of citing WO application

IPC	Total patent references	Type of reference			
		NX	NY	IX	IY
4-digit	87 (69.6%)	56 (70.0%)	1 (50.0%)	25 (71.4%)	24 (82.8%)
7-digit	66 (52.8%)	44 (55.0%)	0 (0.0%)	18 (51.4%)	17 (56.6%)
full	36 (28.8%)	28 (35.0%)	0 (0.0%)	11 (31.4%)	5 (17.2%)
Total	125 (100%)	80 (100%)	2 (100%)	35 (100%)	29 (100%)