

Corporate Science in the Patent System: An Analysis of the Semiconductor Technology*

Antonio Della Malva^a and Katrin Hussinger^{a,b,c}

*^aMaastricht University, School of Business and Economics,
Department of Organization & Strategy (The Netherlands)*

^bK.U. Leuven, Dept. of Managerial Economics, Strategy and Innovation (Belgium)

^cCentre for European Economic Research (ZEW), Mannheim (Germany)

November 2010

Abstract

Corporate scientific publications are often presented as a strategic means for firms to create prior art with the objective to prevent others from patenting related inventions. This paper provides a first analysis of the effects of scientific publications in the patent system. With focus on the semiconductor technology we show that scientific publications by corporations challenge the novelty and the inventive step of patent applications at the European Patent Office (EPO) significantly more than other pieces of prior art. Scientific publications from the public sector, in contrast, exhibit no extra novelty challenging effect. Detailed information from the EPO patent examination procedure allows us to show that corporate publications threaten the novelty of patent applications in combination with other pieces of prior art (rather than as standalone documents). We conclude that corporate scientific publishing can be an effective means for firms to protect their freedom to operate in the complex semiconductor technology if used as a complementary part of a firms' overall IP protection strategy.

Keywords: defensive corporate publishing; freedom to operate; patents; semiconductor

JEL-Classification: O34, O32

* We thank Michele Pezzoni granting us access to his text search algorithm and subsequent “technical” support, Fleur van Dulleman for her precious assistance during the data gathering process and Thorsten Doherr for granting us access to his search engine as well as for providing us with an update of the patent application outcomes. For helpful comments we thank Dirk Czarnitzki, Bronwyn Hall, Cedric Schneider and Brian Wright as well as the participants of the CISS summer school in Turkey, 2010, the Brown Bag Seminar at the KU Leuven, 2010, the EPIP Conference, 2010, and the Innovation Seminar at the UC Berkeley, 2010.

1. Introduction

Innovation can be quintessential for achieving a competitive advantage vis-à-vis competitors and for long-term profitability. The extent to which firms profit from innovation hinges upon their ability to appropriate innovation rents (Cohen et al., 2000, Ceccagnoli, 2009). Traditionally, the effectiveness of appropriation strategies has been evaluated with regard to their success in restraining third parties from using the invention (Levin et al., 1987, Arundel et al., 1995, Cohen et al., 2000, Winter, 2006). A more general view establishes that effective appropriation mechanisms should guarantee the inventor the freedom to operate her technology in the first place (Henkel and Prangerl, 2008). The exclusion of third parties, e.g. via patents or secrecy, is only one way to achieve this. Alternatively, firms can submit technical information about an invention into the public domain to increase the pool of prior art and thereby preclude rivals from patenting related inventions (Parchomocsky, 2000, Johnson, 2004, Adams and Henson-Apollonio, 2002, Merges, 2004, Maurer, 2002, Borchardt, 2007, Buxbaum, 2001, Colson, 2001).¹ This strategy is referred to as defensive or pre-emptive publishing.

The academic literature has developed different answers to the question why firms engage in defensive publishing at the costs of forgoing the right to exclude third parties from using the invention. Economic scholars have analyzed defensive publishing in the context of patent race models (Parchomovsky, 2000, Baker and Mezzetti, 2005, Bar, 2006, Lichtmann et al., 2000, de Fraja, 1993). Patent race models typically consider two inventors that compete for being first to complete and patent a specific invention. The time after which the invention occurs is ex-ante unknown and depends largely on the inventors' investment in research and development. In these models, the loser typically receives no or little return to her investment, while the winner gets all (e.g. Reinganum, 1982). Defensive publishing is

¹ Merges (2004) and Maurer (2002) discuss defensive publishing in the context of genomics (see also Eisenberg, 2000). Examples of firms being highly active in publishing technical advances are IBM, Philips, Hitachi, ICI, Ciba, Siemens, Sandoz, Roche, Hoechst and Toshiba (Hicks, 1995). In the technology field of computer science, IBM and AT&T each published more than Stanford and MIT together in the period 1991-2001. In material sciences, IBM is ranked number two after Tohoku University (Science Watch, 2001).

introduced into these models by allowing for the option to disclose knowledge outside the patent system. In such a scenario, the laggard has incentives to disclose her findings before the leader applies for a patent in order to expand the state of the art and, therewith, increase the patentability requirements for the rival and prolong the patent race.

The patent race explanation for defensive publishing has raised criticisms on both legal and strategic grounds (Eisenberg, 2000). With regard to the U.S. patent legislation, Eisenberg (2000) argues that the first-to-invent rule² and the statutory bars of the U.S. patent doctrine³, which define prior art as of one year prior to the application date, limit the effectiveness of preemptive publications. From a strategic point of view, Eisenberg (2000) criticizes that the disclosure of relevant information outside the patent system might not be beneficial for the laggard since it can speed up the completion of the rival's invention by means of knowledge spillovers (De Fraja, 1993) and invite additional competitors, while at the same time, it increases the patenting threshold for the laggard firm as well (Rinner, 2003).

Scholars in the fields of law and management science suggest defensive publishing as a complementary strategy to patenting (Colson, 2001, Buxbaum, 2001, Rinner, 2003, Johnson, 2004, Henkel and Pangerls, 2008). In this setting, companies ponder the costs and benefits of patent protection for their inventions. Johnson (2004) develops a theoretical framework in which companies can choose between patenting, publishing and keeping the invention secret depending on the respective costs and benefits. A major difference between this model and the patent race models is that the latter models consider defensive publishing as a strategy for the laggard who is unable to patent, while Johnson (2004) models defensive publishing as an equivalent alternative to patenting and secrecy. His model predicts that publishing is most attractive for inventions that are less technically challenging or that can be easily invented around. Publishing can be a defense against "patent pirates" that strive for patent rights with the aim to exploit these patents in litigations. These predictions received partial support from

² The first-to-invent rule allows the patent applicant to prove the invention was completed earlier than the conflicting prior art (U.S. patent doctrine, section 102(a)). This practice is referred as "swearing behind the reference" or "Rule 131 affidavit" (Eisenberg, 2000: 2360)

³ See U.S. patent doctrine, section 102(b).

law scholars and intellectual property (IP) experts who strongly advice to rely on defensive publishing only for minor improvements on existing core technologies which are already well protected (Buxbaum, 2001, Colson, 2001, Rinner, 2003). Henkel and Pangerls (2008) confirm the use of defensive publishing as a complementary strategy to patenting based on interviews with 44 patent experts from 37 publicly listed companies in the German business sector. They strongly dismiss the patent race model: companies would barely know if they are in a race and they would, in particular, not be able to determine their position in the race (whether they are the laggard or the leader).

This study makes two contributions to the literature on corporate disclosure. First, we provide the first systematic (large-scale) analysis of the effectiveness of corporate scientific publications in the patent system. We show that corporate publications can help securing the freedom to operate a technology by blocking the granting of related patent applications as pieces of prior art. Prior empirical evidence on the effectiveness of defensive publishing is limited to case illustrations within the field of genomics (Eisenberg, 2000, Maurer, 2002, Merges, 2004) and interviews with IP experts in the private sector (Henkel and Pangerl, 2008). Second, we show that corporate publications are especially effective against patents in combination with other pieces of prior art rather than as standalone documents. In line with the law and management science literature, this suggests that corporate publications when used for defensive publishing purpose are most effective as a complementary part of a firm's overall IP protection strategy.

Our analysis is based on a sample of semiconductor patent applications in the period 1995-2000 at the European Patent Office (EPO). At least two characteristics render this technology attractive for our analysis. First, it is characterized by a strong reliance upon scientific discoveries (Breschi and Catalini, 2010). Second, the semiconductor technology is a complex technology where a single product relies on several patents. Firms in complex technologies own fragmented patent portfolios and pursue "portfolio maximizing" IP protection strategies in order to avoid being held up by other firms (Hall and Ziedonis, 2001, Ziedonis, 2004).

Semiconductor firms, accordingly, engage in patent portfolio races rather than in traditional patent races (Hall and Ziedonis, 2004). In this context, the dissemination of scientific results by firms can represent an effective (complementary) means to secure their freedom to operate technologies.

Focusing on the European patent system has several advantages too. First, Eisenberg's (2000) critique on the effectiveness of publishing based on legal grounds does not apply since, for one, the EPO has a first-to-file rule in place rather than a first-to-invent regulation, and second, the EPO defines prior art as all inventions being known before the first date of filing. This renders defensive publishing a potentially more effective strategy in the European context than in the U.S. Second, the EPO publishes detailed information on the patent examination procedure including the purpose of individual references to prior art in patent applications (Webb et al., 2005, Harhoff et al., 2005). Based on this information we can distinguish between references that challenge the novelty of a patent application in question and references defining the non-infringing state of the art in a technology field. Further, this information allows us to differentiate prior art that threatens the novelty of a patent application in combination with other pieces of prior art from infringing standalone documents.

The remainder of the paper is organized as follows: section 2 describes the patentability criteria at the EPO and details how scientific publications can impact patent grant decisions. Section 3 describes the construction of the data set and presents descriptive statistics. The fourth section shows the empirical results. Section 5 provides additional tests to strengthen our interpretation of the main findings. The last section concludes.

2. The European Patent System

This section outlines the patentability requirements at the EPO and details how scientific publications can impact patent grant decisions.

2.1 Patentability Requirements

According to article 52(1) of the European Patent Convention (EPC), a patent application has to fulfill three criteria to be granted: a) novelty, b) inventive step and c) industrial applicability. Scientific publications can play a significant role when a patent application's novelty and its inventive step are evaluated. The inventive step demands that the invention "having regards to the state of the art, [it] is not obvious to a person skilled in the art" (EPC, article 56). Article 54(1) of the EPC defines novelty as not being part of the state of art. The state of the art is defined as "everything made available to the public by means of a written or oral description, by use, or in any other way, before the date of filing of the European patent application." (EPC, article 54(2)).

Scientific publications can shape the state of the art in technology fields and by doing so they can push forward the novelty frontier. Consequently, scientific contributions can affect the requirements for the inventive step as defined by the current state of the art (Barrett, 2002, Bar-Gill and Parchomovsky, 2003, Henkel and Pangerls, 2008, Parchomovsky, 2000).

2.2. Searches for Prior Art

Once a patent application is filed, the search divisions of the EPO carry out a patentability search. The aim of such searches is to collect technical information which defines the state of the art relevant to judge the appropriateness of the scope of the legal protection as requested in the application document.⁴ Examiners base their searches on the exact wording of the claims. The result of their search is documented in the so-called search report which lists all references to previous technical documentations in order to determine the novelty and the inventive step of the patent application at the time of the filing.⁵ The search of relevant prior

⁴ At the United States Patents and Trademarks Office (USPTO), in contrast, patent applicants are subject to the "duty of candor". This means that they have to submit a list of all relevant prior art along with the patent application. Non-compliance can result in a rejection of the patent application. Given that the duty of compliance the applicants, patent applicants at USPTO provide a disproportionate high number of references (Silverman, 2003).

⁵ The evaluation of the third criterion of industrial applicability usually does not imply further assessment based on prior art (Michel and Bettels, 2001).

art follows the *Guidelines for Examination in the European Patent Office*⁶ in order to guarantee a certain quality standard as well as equal treatment of all EPO patent applications.

For this study, at least two features of the EPO search report deserve attention. Firstly, the examination guidelines explicitly require examiners to be objective and selective when defining the documents referred to as prior art. In most of the cases one to two documents are sufficient to determine the scope of the patent application in question (Michel and Bettels, 2001). This parsimonious and objective approach ensures that the references are a relevant subsample of the actual state of the art rather than an overview on the subject-matter of the invention (Harhoff and Reitzig, 2004, Harhoff et al., 2005).

A second feature of the EPO search reports is the classification of references according to their relevance in terms of the three criteria of patentability (Dernis et al., 2005, Webb et al., 2005, Harhoff et al., 2005). Two types of citations challenge the patentability of patent applications (citation types X and Y). X-type citations are documents showing essential features of the invention under investigation or at least questioning the novelty and the inventive step of these features if taken alone. Y-type citations question the inventive step claimed in the invention being examined, when combined with other pieces of prior art.⁷

The minimalistic approach to prior art searches and the categorization of the references according to the patentability criteria enable us to assess the impact of scientific contributions in the process of patent granting. The principle of parsimony regarding the references allows us to identify the scientific “shoulders” on which new technical improvements stand. The categorization of references tells us to which extent the novelty and inventive step of patent applications is threatened by prior art and by corporate science, in particular.

⁶ See <http://www.epo.org/patents/law/legal-texts/guidelines.html>.

⁷ Guellec et al. (2009) show that X- and Y-type references to the patent literature significantly hinder positive patent grant decisions. X- and Y-type references have been used to identify weak patents when predicting the likelihood of patent opposition (Harhoff and Reitzig, 2004, Reitzig, 2004, Czarnitzki et al., 2009). Further, patents receiving X and Y citations by future patent applications have been shown to increase firms’ market value (Czarnitzki et al., 2010) as well as firms’ value in the market for corporate control (Grimpe and Hussinger, 2008a,b).

3. Data Collection and Descriptive Statistics

3.1 Patent Data Sources

Our primary source of information is the EPO database, which includes information about all patent applications since its foundation in 1977. From this database we retrieve all patents assigned to the semiconductor technology in the years 1995-2000.⁸ We define the technology field of semiconductor according to the Observatoire des Sciences et des Techniques (OST-INPI/FhG-ISI), also known as the Fraunhofer technology classification (Schmoch, 2008). This categorization aggregates the narrowly defined patent technology classes of the International Patent Classification (IPCs) to technology fields. Semiconductor applications are identified by the 4-digit IPC code “H01L”. After removal of some incomplete filings, we identified 8,385 different semiconductor patent applications in our time period of interest. For these patents we retrieved the application year, IPC classes, inventor and applicant information as well as the patent decision outcome as of 2009 (granted, rejected, pending, withdrawn).

Information about the references made by our patents were gathered from the “EPO/OECD patent citation database”. This database covers all types of references to prior art including references to patents and references to the non-patent literature (NPL) (Webb et al., 2005). The 8,385 semiconductor patent applications cite 37,307 other patents and make 9,925 references to the non-patent literature. We also extracted the classification of references from the search reports from this data source to be able to distinguish between blocking references (X- and Y-type of references) and such that describe the non-infringing state of the art in a particular field of semiconductor technology. Furthermore, we extracted information on patent priority claims from these files. A patent application may claim priority from a previously filed application at a different patent office in order to take advantage of the filing

⁸ Later patent applications are not considered because we are interested in the outcome of the patent application decision which can take place several years after the application date. Earlier patents could not be considered as we have only limited information on (cited) scientific publications issued before 1988.

date of that earlier application. The pool of prior art is then restricted to documents that were published before that first filing date.

3.2 Construction of the Sample

While references to patent documents can be easily handled since they are identified by their patent application number this is not the case for the cited non-patent literature. NPLs are documented in form of text fields including the article title, author information, journal name and publication year. The format of these text fields differs across the records. This is partly explained by the fact that the NPL records are quite heterogeneous encompassing a range of documents such as scientific articles, books, abstract meetings and abstract services but also trade journals and presentations at trade fairs. Accordingly, not all NPLs represent a link to scientific work (Harhoff et al., 2003a). Across all technology fields, roughly about 65% of the NPL references are of scientific nature (Callaert et al., 2006). For our sample, we manually coded the NPL records in order to retrieve journal and author names. The result of this exercise was a list of all references to the patent and non-patent literature that are cited in our semiconductor patents.

In the next step, we linked the list of NPL references to the SCI-Expanded (Science Citation Index Expanded) file of the Web of Science publication database in order to identify NPLs that correspond to scientific journal publications. The SCI-Expanded database provides access to articles published in the most important 8,130 journals across different disciplines. In total, 80% of all scientific journals are covered by this database (Van Vianen et al., 1990). NPLs were matched with the publication file based on author names, journal names and publication years. A computer-supported search algorithm was used to facilitate the link. Each potential match proposed by the algorithm was checked manually. The link allows us to distinguish references to scientific journals from other NPLs. We recovered 2,124 scientific contributions being cited in 893 patent applications in our sample.⁹ Furthermore, we extracted the author affiliations from the publication files in order to be able to distinguish corporate

⁹ The majority of NPLs that did not represent links to scientific journals are references to specialized disclosure platforms, like IP.com, and company bulletins, like the IBM bulletin.

publications from scientific advances published by the public sector.¹⁰ The top publishing companies in our sample with more than 20 articles cited in the patent literature are AT&T, IBM, NTT, Hitachi, NEC, Phillips, Siemens/Infineon, Motorola, Matsushita, Toshiba, Texas Instruments and Fujitsu.

The final sample is subject to some restrictions: First, we only kept granted and rejected patent applications. Pending patent applications were excluded since we have to observe the final decision on a patent application to be able to evaluate the impact of corporate scientific publications on the patent examination outcome. Withdrawn patent applications were excluded as well since withdrawals present a patent application outcome which is outside the control of the patent office. In total, 1,812 pending or withdrawn patent applications were dropped. Second, we restrict our analysis to corporate patent applications as there is some evidence that patent applications from scientific institutions or academic inventors are significantly different regarding the nature of the protected invention (Trajtenberg et al., 1997, Czarnitzki et al., 2009). In this step, 347 patent applications without any firm participation were excluded. In total, our final sample consists of 6,226 rejected or granted corporate patent applications in the field of semiconductor technology. Table 7 in the Appendix shows the distribution of our patent applications across different disciplines within semiconductor technology.

We created a unique identifier for patent applicants in our sample to account for unobserved effects on the level of the patent applicant in the empirical analysis. As before, a text-based search algorithm was used. Each potential “group” of patent applicants referring to the same corporation proposed by the algorithm was manually checked. Multiple patent applicants on patent applications were treated as research consortia and, hence, received their own id as a group. The top patenting firms in our sample are Infineon/Siemens,

¹⁰ We use that term “public sector” in the broadest sense in this paper. Publications from the public sector include all non-corporate publications such as scientific publications taken out by universities, public research institutions, governmental institutions, private persons etc.

STMicroelectronics, NEC, Applied Materials Inc., Sumitomo, Texas Instruments, Philips, Canon, Matsushita, and IBM with more than 100 granted or rejected patent applications.

A limitation of the data gathering process was that the SCI-Expanded database is only available to us from 1988 onwards. Hence, we could not classify NPLs that were published earlier as being scientific or not. Further, we were not able to identify some references in the SCI-Expanded database due to incomplete or ambiguous information in the patent reference list. We account for both in the empirical analysis by including a first dummy variable indicating that a patent application cites NPLs published before 1988 and a second dummy variable indicating citations to NPLs with unidentified nature in our regression model. In total, there are only 4% of the all NPLs for which we could not retrieve all necessary information.

3.3 Descriptive Statistics

Table 1 presents descriptive statistics for the variables used in the later multivariate regression models for the patent application outcome.¹¹ The descriptive statistics show that granted patent applications differ from rejected ones in several ways. Granted patents, for instance, make more references to prior art (*REF*). The number of backward citations has been given different interpretations in the previous literature. Some scholars argue that patents with many backward citations are of a more incremental nature and that they target more crowded technology fields in the U.S. context (Lanjouw and Schankerman, 2001). Others claim that patents with many backward citations protect an invention with a large scope that invites the patent examiner at the EPO to delineate the patent claims by inserting more references to relevant prior art (Harhoff et al., 2003a). As a measure for the scope of patents backwards citations are an important variable in our empirical analysis since novelty challenging (X- and Y-type) citations are given by the examiner to limit the scope of patent applications. In the empirical specification, we use the logarithm of the total references to account for the skewness of this variable's distribution. As a second measure for the scope of

¹¹ Table 8 in the Appendix shows bivariate correlation for our main regressors.

a patent application the number of 4-digit IPC classes is used (Lerner, 1994).¹² Due to the skewness of the this variable's distribution the logarithm is used. If this variable is larger than zero the patent application is attributed to other technology fields than semiconductor. The correlation between both patent scope proxies is low with 20% (see Table 8 in the Appendix).

Focusing on the novelty challenging references to prior art, Table 1 shows, that, not surprisingly, granted patent applications receive fewer blocking references (*XY REF*) than rejected ones. Note that a patent application can still be granted if it receives many references classified as X or Y. This can be the case for patent applications with many claims or with a broad scope. X- and Y-type references may pertain to single claims and the remaining claims can be strong enough to get a (modified) application granted. Since the number of total references as a proxy for the scope of the patent application and the number of blocking references are correlated by definition, we normalize the number of blocking references by the number of total references ($XY\ REF/REF$) for the specification of the empirical model.

Considering the total number of references to the non-patent literature (*NPL*), Table 1 shows that rejected patent applications make significantly more references to the non-patent literature. Rejected patent applications make significantly more reference to the scientific literature (*JOUR*) as a subgroup of citations to the non-patent literature as well. The average patent application in semiconductor cites .26 scientific sources. This underlines the science-basedness of semiconductor technologies. The fact that the number of scientific NPLs is higher among rejected patent applications than among granted ones suggests that, in the field of semiconductor, science does not only represent prior art that lays a foundation for technology development, but that scientific publications also challenge the novelty of patent applications.

The variable $XY\ JOUR\ REF/XY\ REF$ presents the average share of scientific references among the blocking type of references the patent applications receive. An average of 5% of the blocking references per patent application stem from scientific publications. Again, the

¹² The number of claims is unfortunately not available to us.

share is significantly higher among rejected patent applications. Distinguishing between X- and Y-type of references among journal references shows that 3% of the blocking type references per patent application refer to scientific publications that are blocking as standalone documents (*X JOUR REF/XY REF*) while another 2% refer to Y-type citations to science that is blocking in combination with other pieces of prior art (*Y JOUR REF/XY REF*).

Regarding the origin of the novelty threatening scientific publications, an average of 4% of the blocking references to science per patent refers to corporate publications (*XY comp JOUR REF/XY JOUR REF*). On average, 4% of these refer to X-type citations to corporate scientific publications (*X comp JOUR REF/XY JOUR REF*) and 2% refer to Y-type references to science taken out by the private sector (*Y comp REF/XY JOUR REF*).

In addition to our main variables of interest and the control variables for patent scope, we have a number of additional control variables. We control for the time that passed between the patent application at EPO and the final grant decision. More complex technologies presumably need more time to be evaluated (Sampat et al., 2003). The logarithm of this variable as measured in years (plus one) is used to account for the skewness of its distribution. Further, the (log of the) number of inventors is used as a proxy for the size of the innovation project. The major share of investments into innovation projects typically represents the wages of the R&D personnel. The number of inventors can, hence, be seen as a proxy for the investment made for a patent application.

Also, we control for the priority of the patent applications in our sample. A patent examiner might use the priority application of the EPO patent application as a reference when searching for prior art. We use four dummies to control for priority at different patent offices, namely the USPTO (2,220 patent applications), the Japanese Patent Office (2,271), the EPO

(473) and national European patent offices (1,561). Patents with priority at other patent offices worldwide are used as the benchmark (257).¹³

Moreover, a dummy variable is used to indicate if a patent application has co-applicants from the public sector (*PUBLIC*). In total, 91 granted and rejected patent applications in our sample are applied for by private-public research cooperations. A set of application year dummies is used to control for time trends because patent pendencies at the EPO increased sharply since the later 1980s (Harhoff and Wagner, 2009). In order to control for technology subclasses within semiconductor eight subfield dummies are used (see Table 7 in the Appendix for the definition of these dummy variables).

Lastly, we use two dummy variables to account for weaknesses of our database. First, a dummy variable controls for patent applications with NPLs published before 1988 as we could not retrieve information about the authors and author affiliation of these publications from the Web of Science. A second dummy variable indicates references to NPLs that could not be classified as being scientific or not due to incomplete or ambiguous information on their citations records. In total, for 8% of the patents we have incomplete information.

Table 1 about here

4. Empirical Results

In this section, we estimate probit models to analyze the effect of scientific publications on the patent grant decision. We model the patent grant decision d_i as a function of the different types of prior art cited in the search report for the focal patent application i . d_i takes the value one if patent application i gets granted and zero is if the examination at EPO leads to a rejection. Due to the high correlation of the different types of references to prior art we use a cascade specification, i.e. we orthogonalize each patent reference variable by the next more general group of references:

¹³ We do not include applicant countries in our empirical specification since they are highly correlated with the priority dummies. Most applicants get priority at the patent office of the country in which they are located.

$$d_i = \beta_1 \log(REF_i) + \beta_2 \frac{XY REF_i}{REF_i} + \beta_3 \frac{XY JOUR REF_i}{XY REF_i} + \beta_4 \frac{XY comp JOUR REF_i}{XY JOUR REF_i} + \beta_5 X_i + u_i$$

The interpretation of the coefficients of this specification is that each ratio presents a premium (or a discount) on the former one. For instance, if the total number of references (REF) would have a positive impact on the patent grant decision and the share of novelty challenging cites among them ($XY REF/REF$) would have a negative impact, the estimated coefficient for the latter variable would reflect a discount on the positive effect of the total number of references. This logic applies to all variables in the cascade. The share of scientific publications among the references with a blocking potential is depicted by $XY JOUR REF/XY REF$ and indicates whether scientific prior art among the blocking references is more (or less) relevant for the patent grant decision than the average novelty challenging citation ($XY REF/REF$). The last variable of central interest ($XY comp JOUR REF/XY JOUR REF$) presents the share of company publications among the challenging scientific references. This variable identifies whether corporate science is more (or less) challenging for the novelty of patent applications than the average blocking reference to science ($XY JOUR REF/XY REF$). Finally, X_i presents the vector of control variables as described in the previous section.

The standard probit model relies on the assumption of homoscedasticity of the error term u_i . We estimate heteroscedastic probit models since a violation of the homoscedasticity assumption leads not only to biased standard errors, but also to inconsistent coefficient estimates (see Greene, 2000). We replace the homoscedastic variance σ by $\sigma_i = \exp(\omega_i' \alpha)$, where ω_i represents the regressors determining the variance function and α the additional coefficients to be estimated.¹⁴ ω_i includes five application year dummies, eight technology field dummies and the (log of) the time until a decision about a patent application is reached. We tested for the presence of heteroscedasticity in our probit models by means of LM tests (see Greene, 2000). The test statistics presented at the bottom of Table 2 show that homoscedasticity is rejected in all cases. Hence, the coefficients of standard probit models

¹⁴ Note that robust standard errors do not help if heteroscedasticity is found in nonlinear models.

would be misleading. In addition, we use clustered standard errors on the level of the applicant to account for the fact that the error terms of the empirical model might be correlated for patent applications by the same applicant.

The estimation results summarized in Table 2 show several interesting findings. Concerning the baseline specification (model 1), which includes only the control variables, we find, for instance, that patents citing more prior art are more likely to become granted. This result does not support the interpretation that patents with many backward citations are of more incremental nature (Lanjouw and Schankerman, 2001). It is rather in line with previous research arguing that backward citations reflect the scope of the patented technology (Harhoff et al., 2003a). The finding that patents with references to patents assigned to many different technology classes is more likely to become granted points in the same direction. Patents associated with many different technologies presumably have a greater patent scope, which correlated with the importance of patents (Lerner, 1994) and renders a positive granting decision more likely.

The next control variable captures the effect of time since application. The longer the patent examination takes the more likely is a grant. This suggests that patent applications which require a longer examination time are more complex (Sampat et al., 2003) and potentially more novel. The number of inventors as a proxy for the size of the innovation project has a positive, but weakly significant effect. There is no difference between purely private patents and patents taken out in alliance with the public sector regarding the likelihood of a positive grant decision. Lastly, application year dummies, priority dummies and technology class dummies are jointly significant as LR-tests at the bottom of Table 2 show. In particular, the priority dummies are highly significant, which is mainly driven by the fact that patents with a U.S. priority are significantly less likely to become granted at the EPO. The estimated effects of the control variables are robust across the different specifications presented in Table 2.

In model 2, we add the share of all references (patents and scientific and non-scientific NPLs) which are classified as prior art that threatens the novelty of patent applications ($XY\ REF/REF$) as an additional regressor to our specification. These references are made to limit the scope of patent applications by evaluating their novelty with regard to prior inventions. As expected, the estimated coefficient clearly indicates that these references that patent examiner includes to delineate the patent claims make a positive grant decision less likely (Guellec et al., 2009). The marginal effect equals -0.14 at the means of all other variables indicating that the likelihood of a positive grant decision decreases by 14% in response to a unit change of the share of blocking references, i.e. if a patent application without any blocking reference would suddenly have only blocking references. This maybe low effect of blocking references shows that a patent application that receives many blocking references can well be granted if claims are removed or modified and the scope of the patent application is narrowed.

Model 3 adds the share of scientific publications that are classified as novelty threatening ($XY\ JOUR\ REF/XY\ REF$) to the specification. The results show that novelty threatening journal publications increase the probability of getting a patent application rejected significantly. Accordingly, in the field of semiconductor, scientific NPLs do not only reflect scientific results that lead to technology development. In this science-based technology field, scientific publications have a power to threaten the novelty of patent applications on top of the blocking power of an av. The regression results suggest that scientific articles even have an extra blocking effect on top of the average blocking citation which is captured by the variable ($XY\ REF/REF$). The marginal effect equals -0.17 at the means of all other variables indicating that the likelihood of a positive grant decision decreases by 17% in response to a unit change of the share of scientific blocking references among all blocking references, i.e. if a patent application without any scientific blocking reference would suddenly have only scientific blocking references.

Model 4 investigates whether corporate scientific publications ($XY\ comp\ JOUR\ REF/XY\ JOUR\ REF$) are responsible for this extra blocking effect of scientific NPLs. The results show

that the extra blocking effect of science in the patent system stems from company publications. If company publications are taken into account separately on top of the overall effect of blocking scientific NPLs the latter variable becomes insignificant. This suggests that only corporate publications have a blocking effect on patent applications on top of the overall blocking effect of X- and Y-type citations. Scientific publications taken out by the public sector are to the same extent blocking as the average piece of prior art classified as X or Y. This result is in line with the presumption that there is a strategic consideration behind company publications. The marginal effect of the blocking company citations equals -0.11 at the means of all other variables suggesting that the likelihood of a positive grant decision decreases by 11% in response to a unit change of the share of company references among all blocking scientific references, i.e. if a patent application without any scientific blocking reference to company publications would suddenly have only scientific blocking references to the private sector.

Model 5 distinguishes between X- and Y-type citations to corporate publications and to scientific references in general. A very interesting result appears. Only the share of Y-type citations to the private sector has an additional blocking effect (on top of the average cited blocking document). This suggests that corporate publications are only effective if combined with other pieces of prior art and not as standalone documents (X-type citation). The marginal effect equals -0.16 at the means of all other variables suggesting that the likelihood of a positive grant decision decreases by 16% in response to a unit change of the share of Y-type company references among all blocking scientific references.¹⁵ The finding that company publications have a blocking power if combined with other pieces of prior art (and not as standalone documents) is in line with the predictions from the law and management science literature that publishing can be a successful IP protection strategy if it is combined with stronger IP protection means, like patents.

Table 2 about here

¹⁵ The estimated marginal effects for the other significant regressors do not increase significantly in size if more variables are added.

Table 9 in the Appendix shows that our results are robust if we delete the most active patent applicants with more than 200 patent applications each. We further checked whether the journal publications that are cited as blocking prior art are citations to the companies' own scientific publications. If the fraction of the blocking citations to science taken out by the same applicant is added as a separate variable to model 4 the estimated effects do not change significantly and the effect of journal self-citations is not statistically significant.

5. Are Corporate Publications More Applied?

The results presented in the previous section have shown that corporate science has a superior power to block patent grant decisions at the EPO than the average reference marked as blocking by the patent examiner; this effect does not hold for science produced in the public sector. This can be the outcome of strategic disclosure by companies to secure the freedom to operate a technology by blocking patent applications for related technologies. An alternative explanation could be that corporate publications address more applied puzzles than scientific publications from the public sector so that they are more relevant for the evaluation of the novelty of patent applications.

The differences between science produced in the private and public sector are largely documented (e.g. Dasgupta and David, 1994, Stokes, 1997, Fleming and Sorensen, 2004). Shortly, the curiosity-driven public scientist strives for intellectual challenges, which might not be of any industrial interest. Science has priority over commercialization and the incentives for innovators are significantly determined by peer recognition and career awards such as tenure rather than by monetary rewards (Merton, 1973). The corporate scientist, in contrast, is hired for the explicit purpose of generating knowledge with a commercialization potential for the firm (Cohen and Levinthal, 1990). These different conditions and motivations for corporate and public science cause that science in the private sector is typically more applied than research conducted in the public sector (Dasgupta and David, 1994, Trajtenberg et al., 1997, Czarnitzki et al., 2009). If the difference in the nature of incentives and goals in both sectors was reflected in scientific publications (cited by our

semiconductor patents), the more applied character of the corporate publications could explain the superior blocking power in the patent system. In this section, we therefore investigate whether there is evidence for corporate publications being more applied than science published by the public sector. Several tests are presented.

As a first check we compare the scientific journals in which the corporate and public articles that are cited in our patents are published. If corporate science was more applied, we would expect that companies publish in more applied journals. Table 3 presents a list of the journals in which the majority of our articles cited in patents are published. It appears that all of the journals that are cited in patents are applied. There are no systematic differences regarding the distribution of corporate and public sector publications across journals. A Mann-Whitney-Wilcoxon rank-sum test on the rankings of publication frequency across journals are equal at their medians (p-value = 0.66).¹⁶

Table 3 about here

Previous literature has established that university patents take longer to get granted and are cited later as compared to corporate patents (Sampat et al., 2003). It is argued that this is because university patents cover more basic technologies and are, hence, more difficult to be evaluated by the patent office and adapted by other inventors. Following this logic, we can extend the same rationale for scientific publications and propose that publications from the public sector should be cited later than corporate publications if they address more complex scientific advances.

Proposition 1: If corporate science cited by patent applications is more applied than science taken out in the public domain, it should be cited earlier.

We focus on the time lags from the publication year of the scientific articles cited in our patents and the year in which they receive the first citation in a patent application. Table 4

¹⁶ We also ran two-paired t-tests that confirmed the previous finding. For a matter of robustness, we ran the same test for the distribution of shares in the first 20, 30 and 50 journals in the total sample. The results indicate no differences in the distributions (p-value = 0.94 for the first 20 journals; p-value = 0.71 for the first 30 journals; and p-value = 0.43 for the first 50 journals).

shows that company publications are cited, on average, six months later than those taken out by scientific institutions. A t-test rejects the null-hypothesis that there is no difference between the average citation lags for corporate publications and publications taken out by the non-profit sector at the 1% level of statistical significance. This suggests that – conditional on being cited in patents at all – public science is cited earlier than corporate science in the patent literature. If we applied the standard interpretation to the citation lag measure this would mean that corporate science cited in patents is more basic and complex than public science cited in patents since it takes longer until the corporate results are picked up in follow-up inventions.

Table 4 about here

Our third test for potential differences between publications from science and industry is based on citation categories. We submit that if corporate publications reflect scientific advances which are closer to industrial applications they should be more likely to threaten the novelty of patent applications. Scientific publications from the public sector, in contrast, should be more likely to be cited as non-infringing state of the art if they reflect more basic research. Hence, our second proposition reads:

Proposition 2: If corporate science cited by patent applications is more applied than science from the public sector, it should be more likely to be cited as blocking prior art while publications from the public sector should be rather cited as non-infringing state of the art.

We conduct two different tests: first, we distinguish between patent citations to publications that are quoted as blocking (X- and Y-types) and other types of citations, and, second, we use a more disaggregated level of categories distinguishing between X-type, Y-type and the residual group of citations. The latter refer to prior art that explicitly refer to the non-infringing state of art in a technology domain (Webb et al., 2005, Harhoff et al., 2005). If publications from the public sector were more basic, we would expect to find them primarily in this citation category. Pearson X^2 -tests show that the distribution of publications from the public sector across citation type categories is not significantly different from the citation type

distribution of corporate publications (see Table 5). Hence, there is no evidence for corporate research being more challenging with regard to the novelty criteria for patent applications than public science.

Table 5 about here

The tests above focused on the potential differences between scientific publications taken out by the corporate and private sector within our sample of publications that are cited in patent applications. A final test compares these publications to a control group of publications that receive no citations by the patent literature. The aim of this comparison is twofold: first, we want to investigate whether publications cited in patents are more applied than the control group and, hence, less important in science (as visible from the citations these publications receive by other journal articles). Second, we are interested in potential differences between scientific publications taken out in the private and corporate sector regarding their importance in science. If corporate publications are more applied they should receive fewer citations in the scientific literature. Lastly, we investigate whether there is a difference between publications cited in patents with regard to their origin (public versus private sector). We submit three propositions:

Proposition 3a: If science cited in patent documents is more applied than science only contributing to the scientific realm, it should receive fewer citations from scientific articles.

Proposition 3b: If science generated in the public sector is more basic than corporate science, it should be cited more often in scientific journals.

Proposition 3c: If public science cited in patent documents is more basic than corporate science cited in patent documents, it should be cited more often in scientific journals.

In order to conduct a test for our propositions we drew a random sample of 100 publications cited by patents in our sample. Half of them are authored by scientists in the private sector, half of them by scientists working in the public sector. For these publications, we retrieved all scientific articles published in the same journal issue from the Web of Science database as a control group of scientific publications that are not cited in patents. We

then manually classified the affiliation of the authors of the control publications. After having excluded scientific articles without affiliation details and those that are co-authored by scientists in public institutions and corporate laboratories, we are left with a sample of 4,892 scientific articles, 3,702 from scientists working in the public sector and 1,190 from corporate scientists.

We run count data models for the number of citations these articles receive by journal articles (up until May 2010) as a measure for their importance in science. Most important regressors in the model are a dummy that indicates whether the author is a scientist in a non-profit organization (*Public*), a dummy whether the publication is cited in patents (*Cited in patents*) and an interaction term between these two variables. In addition, we include the number of pages per article as a proxy of the amount of findings reported in the document, publication year dummies¹⁷ and journal dummies as additional regressors.

Table 6 shows the empirical findings. The estimation results show no significant difference in the number of citations by scientific journals with regard to the author affiliation (private versus public sector). Hence, we do not find support for the hypothesis that corporate articles are less important in the scientific domain than articles taken out by the non-profit domain. An interesting result is, however, that publications that receive citations by patents are also more frequently cited in scientific journals. This signals that publications which are important in the scientific domain are also those that trigger future advances in a science-intense technology as semiconductor. There is no significant effect of the interaction term of being cited in patents and being authored by a scientist in the public sector.

Table 6 about here

To summarize our test results, we do not find any evidence that corporate publications are more applied than publications by the public sector conditional on being cited in the patent literature: corporate publications are published in the same journals, they are not cited later in

¹⁷ The publication year dummies should account for the truncation of the dependent variable. Note that truncation is a minor problem in this application since the control group is taken from the same publication year as the publications cited in patents.

the patent literature and they are not differently distributed across citations categories. This suggests that there is no difference in nature of the published technologies cited in patents that could explain the blocking power of corporate publications. Furthermore, there is no difference between publications from the corporate and public domain regarding their importance in science. These findings underline the possibility of a strategic explanation for our previous finding that corporate publications are more powerful in the patent system. For instance, if corporate articles were strategically published as complements to other pieces of prior art as the empirical results from the previous section proposed.

6. Conclusion

Defensive disclosure of technological knowledge by companies to preempt rivals' patents has attracted the attention of academic scholars in the field of law, economics and management science (e.g. Merges, 2004, Parchomovsky, 2000, Johnson, 2004). The academic discussion about this topic has raised two main streams of thoughts. The first branch, mainly carried out by economists, sees open disclosure of technological advances as a catching up strategy for laggards in patent races. In these models the laggard publishes preliminary results that are not patentable in order to prolong the patent race (Parchomovsky, 2000, Baker and Mezzetti, 2005, Bar, 2006, Lichtmann et al., 2000, De Fraja, 1993). Scholars in law and management science, on the contrary, consider open disclosure as a complementary strategy to patenting (Colson, 2001, Buxbaum, 2001, Rinner, 2003, Johnson, 2004, Henkel and Pangerls, 2008). The latter perspective finds partial confirmation in interviews with patent experts in the corporate sector, while the patent race model is largely rejected by practitioners (Henkel and Prangerl, 2008).

This study provides for the first time systematic (large-scale) evidence on the effectiveness of corporate publications for securing the freedom to operate a technology by blocking related patent application. We focus on the field of semiconductor technology as an example of a science-based, complex technology. Our empirical analysis reveals several interesting findings. First, we show and confirm that the technology field of semiconductor is heavily

based on scientific knowledge. An average patent application in semiconductor cites .26 scientific journal articles. Second, we show that science in semiconductor does not exclusively serve as a foundation for technology development, but that it also challenges the novelty of patent applications. In particular, company publications are effective in blocking related patent applications and, accordingly, have the power to serve as a device for securing the freedom to operate a technology.

In a second step of the analysis, we use detailed information from the patent examination process at the EPO that allows us to show that corporate publications are, in particular, novelty threatening if combined with other pieces of prior art, while they are not effective as stand-alone documents. This suggests that corporate publications in semiconductor are effective in securing the freedom to operate as complement to other pieces of prior art as suggested in the law and management literature (Colson, 2001; Henkel and Prangerl, 2008; Rinner, 2003). This finding is intuitive for a complex industry like semiconductor where firms pursue “portfolio maximizing” IP protection strategies in order to avoid being held up by other firms (Hall and Ziedonis, 2001, Ziedonis, 2004).

We can reject that corporate scientific publications have a superior power to block patent applications because they deal with scientific advances that are more applied and closer to commercial ends than scientific publications produced in the public sector, and are, hence, potentially more relevant for the evaluation of the novelty of patent application. We can rule out this alternative explanation based on a number of tests. The results show that there is no significant difference in the nature and significance of publications by corporation and by non-profit organizations, both in the technology and in the scientific domain.

A limitation of this study is that we cannot provide evidence for the motivation behind corporate scientific publications. In other words, we cannot establish that firms publish scientific advances to intentionally preempt (patent) competition. We acknowledge that there are several other reasons for companies to disclose their inventions (Penin, 2007). Companies can, for instance, disclose technical advances to allow upstream suppliers improving their

products (Harhoff et al., 2003b). Disclosure of technologies can increase the size of downstream markets (Harhoff, 1996) or trigger feedback from users in downstream markets (Penin, 2007). Knowledge can also be disclosed with the intention to implement a technology as a standard (Spencer, 2003) or to achieve compatibility with interdependent products (Cusumano and Gawer, 2002). Scientific publications can further be used as a means to increase firm reputation (Hicks, 1995), to motivate researchers in the private sector (Henderson and Cockburn, 1994) and to attract the best employees (MacMillon and Hamilton, 2000). We want to stress that these alternative motivations for corporate scientific publications do not impact our findings that corporate science has superior power to block patent applications and that this effect stems from the combination of corporate science with other pieces of prior art. Our study just does not allow the interpretation that firms intentionally publish corporate technologies to deter patents on related technological matters.

As it is not possible to investigate the motivations behind corporate publications on a large scale our study has to be seen as a complement to case study evidence on this topic. Interviews with industry expert would be needed to uncover how publishing and patenting is strategically used at the firm level. Information on the firms' technology portfolio and on the different IP protection tools that are used at the level of the individual innovation project (and their combinations) would be necessary which renders this research question impossible for large-scale empirical analyses.¹⁸

¹⁸ An approach applied in prior studies to assign patents and publications to innovation projects is to define patent-publications pairs within a firms' patent and publication portfolio (Ducor, 2000, Murray and Stern, 2007, Lissoni and Montobbio, 2008). Such an approach relies either on the knowledge of a person skilled in the art who is able to define patent-publication pairs by studying their content (Ducor, 2000, Murray, 2002, Murray and Stern, 2007), which renders this method impossible for large samples, or it is based on co-word analyses of patent documents and scientific publications (Lissoni and Montobbio, 2008). In our eyes, it is not evident that the latter approach would yield valid matches in a complex, cumulative technology like semiconductor where one invention is based on many different building blocks that are likely to be relevant for more than one innovation projects. The fact that semiconductor has many big players holding huge patent and publication portfolios would make the definition of individual innovation projects even less credible.

Similarly, cross-citations between patents and publications are not a perfect means for defining patent-publication pairs. The majority of references to prior art in EPO patent applications are added by the patent examiner. This means that publications cited in patents do not necessarily indicate that the cited publications refer to the same innovation project as the patent application. Further, the patent examination guidelines dictate that only the most important and earliest sources are to be cited. Hence,

Furthermore, we only focus on scientific articles as a channel for corporate strategic disclosure. We are aware of the fact that there are alternative platforms for corporate disclosure as, for instance, specialized publication platforms like “Research Disclosure” and “IP.com”. Henkel’s and Pangerl’s (2008) interviews with IP experts in industry, however, underline the importance of scientific publications for defensive publishing. They conclude that companies rate scientific publications as the second most effective mechanism to disclose strategically (after internet providers specialized in public disclosure). In the context of a science-based industry like semiconductor, scientific publications might be even more important for strategic disclosure by companies.

Another limitation is that our results for the semiconductor technology cannot be considered as representative for other technology fields. Semiconductor is one of the most science-based technologies. Most likely, corporate publications are less effective for blocking patent applications in other technological areas. Semiconductor is, furthermore, a complex technology where firms engage in patent portfolio races rather than in traditional patent races (Hall and Ziedonis, 2004). It cannot be taken for granted that our result that publications block patent applications in combination with other pieces of prior art applies to other technology sectors as well. In particular, for discrete industries, where the patent race models are a more attractive theoretical framework to study corporate publications, more research should be needed.

we would miss many citations to publications from the same project. Patent citations in publications might be a more promising venue for identifying patents and publications belonging to the same project. Again, there is no guarantee that all patents belonging to the same project are cited in publications. Further, patents belonging to different projects can be cited in publications. Hence, the question remains how reliable such an approach would be for a complex, cumulative technology with many agents with huge publication and patent portfolios.

References

- Adams, S., and V. Henson-Apollonio (2002). "Defensive publishing: a strategy for maintaining intellectual property as public goods", *International Service for National Agriculture Research*, 53, 1-8.
- Arundel A., G. van de Paal and L. Soete (1995). "Innovation strategies of Europe's largest industrial firms. PACE Report", MERIT, Maastricht, The Netherlands.
- Baker, S., and C. Mezzetti (2005). "Disclosure as a strategy in the patent race", *Journal of Law and Economics*, 48, 173-194.
- Bar, T. (2006). "Defensive publications in an R&D race", *Journal of Economics and Management Strategy*, 15, 229-254.
- Bar-Gill, O., and Parchomovsky, G (2003). "The value of giving away secrets", *Virginia Law Review*, 89, 1857-1895.
- Barrett, B. (2002). "Defensive use of publications in an intellectual property strategy", *Nature Biotechnology*, 20, 191-193.
- Borchardt, J. (2007, May 31). "The Role of Defensive Publications in Intellectual Property Protection", Called on November 24th, 2009, from Lab Manager: http://www.labmanager.com/articles_pf.asp?ID=62
- Breschi S. and C. Catalini (2008). "Tracing the linkages between science and technology: an exploratory analysis of the research networks among scientists and inventors", *Research Policy*, 39(1), 14-26.
- Buxbaum P. (2001). "Publish or perish", *The Bridge*, Fall 2001, 2-8.
- Callaert J., B. Van Looy, A. Verbeek, K. Debackere and B. Thijs (2006). "Traces of prior art: an analysis of non-patent references found in patent documents", *Scientometrics*, 69(1), 3-20.
- Ceccagnoli, M. (2009). "Appropriability, preemption, and firm performance", *Strategic Management Journal*, 30, 81-98.
- Cohen W. and D. Levinthal. (1990). "Absorptive capacity: a new perspective on learning and innovation", *Administrative Science Quarterly*, 35, 128-152.
- 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)", Working Paper no. W7552, National Bureau of Economic Research.
- Colson, T.J. (2001). "Leveraging the power and value of intellectual property rights", *Ivey Business Review*, July/August, 28-33.
- Czarnitzki, D., K. Hussinger and C. Schneider (2009). "Why challenge the ivory tower? New evidence on the basicness of academic patents", *Kyklos* 62, 488-499.
- Czarnitzki, D, K. Hussinger and B. Leten (2010). "Patent Value and Citation Heterogeneity", mimeo.
- Cusumano, M. and A. Gawer (2002). "Platform Leadership. How Microsoft, Intel and Cisco Drive Industry Innovation", Harvard: Harvard Business School Press.
- Dasgupta P. and P. David (1994). "Toward a new economics of science", *Research Policy*, 23, 487-521.
- De Fraja, G. (1993). "Strategic spillovers in patent races", *International Journal of Industrial Organization*, 11, 139-146.

- Ducor, P. (2000). "Coauthorship and Coinventorship", *Science*, 289, 873-875.
- Eisenberg, R. (2000). "The promise and perils of strategic publication to create prior art: a response to Professor Parchomovky", *Michigan Law Review*, 98, 2358-2370.
- Fleming, L. and O. Sorensen (2004). "Science as a Map in Technological Search", *Strategic Management Journal*, 25, 909-928.
- Grimpe, C. and K. Hussinger (2008a). "Pre-empting technology competition through firm acquisitions", *Economics Letters*, 100, 189-191.
- Grimpe, C. and K. Hussinger (2008b). "Building and blocking: the two faces of technology acquisition", *ZEW Discussion Papers 08-042*, Mannheim.
- Greene, W..H. (2000), "Econometric Analysis" (4th ed.), New York.
- Guellec D., C. Martinez and M.P. Zuniga (2009). "Pre-emptive patenting: securing market exclusion and freedom of operation", *OECD Science, Technology and Industry Working Paper No. 2009/08*, Paris.
- Hall B.H. and R.H. Ziedonis (2001). "The determinants of patenting in the U.S. semiconductor industry, 1980-1994", *RAND Journal of Economics*, 32, 101-128.
- Harhoff, D. (1996). "Strategic spillovers and incentives for R&D", *Management Science* 42, 907-925.
- Harhoff, D., J. Henkel and E. von Hippel (2003b). "Profiting from voluntary information spillovers: how users benefit by freely revealing their innovation", *Research Policy* 32, 1753-1769.
- Harhoff, D., Hoisl, K., and C. Webb (2005). "European patent citations. How to count and how to interpret them", *LMU Discussion Papers*. Munich.
- Harhoff, D. and M. Reitzig (2004). "Determinants of opposition against EPO patent grants - the case of biotechnology and pharmaceuticals", *International Journal of Industrial Organization* , 22, 443-480.
- Harhoff, D., F. M. Scherer and K. Vopel (2003a). "Citations, Family Size, Opposition and the Value of Patent Rights - Evidence from Germany", *Research Policy*, 32, 1343-1363.
- Harhoff, D. and S. Wagner (2009). "The Duration of Patent Examination at the European Patent Office", *Management Science*, 55(12), 1969-1984.
- Henderson, R. and I. Cockburn (1994). "Measuring competence? Exploring firm effect in pharmaceutical research", *Strategic Management Journal*, 15, 63-84.
- Henkel, J., and S. Pangerl (2008). "Defensive publishing - an empirical study", *DRUID Working Paper No. 08-04*.
- Hicks, D. (1995). "Published papers, tacit competencies and corporate management of the public/private character of knowledge", *Industrial and Corporate Change*, 4(2), 363-400.
- Johnson, J.P. (2004). "Defensive publishing by a leading firm" Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=606781
- Lanjouw, J. and M. Schankerman (2001). "Characteristics of patent litigation: a window on competition", *RAND Journal of Economics*, 32, 129-151.
- Lerner J. (1994). "The importance of patent scope: An empirical analysis", *RAND Journal of Economics*, 25, 319-333.
- Levin R.C., A.K. Klevorik, R.R. Nelson, S.G. Winter (1987). "Appropriating the returns from industrial R&D", *Brooking Papers on Economic Activity*, 75, 783-820.
- Lichtman, D., S. Baker, S., and K. Kraus (2000). "Strategic disclosure in the patent system", *Vanderbilt Law Review*, 53, 2175-2217.

- Lissoni, F. and F. Montobbio (2008). "Inventorship and authorship in patent-publication Pairs: An enquiry in the economics of scientific credit", CESPRI Working Paper 224, Milan.
- Maurer, S.M. (2002). "Promoting and disseminating knowledge: The public/private interface", Paper Prepared for the U.S. National Research Council's Symposium on the Role of Scientific and Technical Data and Information in the Public Domain, Washington, D.C., September 5-6, 2002. *Information Services and Use*, 22(4), 183-189.
- McMillan, G.S. and R.D. Hamilton (2000). "Corporations need to publish or perish", *Research Technology Management*, 43, 8-10.
- Merges, R. (2004). "A new dynamism in the public domain", *University of Chicago Law Review*, 71, 183-203.
- Merton, R. K. (1973). "The Sociology of Science: Theoretical and Empirical Investigation", Edited by N. W. Storer. Chicago: University of Chicago Press.
- Michel, J., and B. Bettels (2001). "Patent citation analysis", *Scientometrics*, 51, 185-201.
- Murray F. (2002). "Innovation as co-evolution of scientific and technological networks: exploring tissue engineering", *Research Policy*, 31, 1389-1403.
- Murray, F. and S. Stern (2007). "Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis", *Journal of Economic Behavior and Organization* 63(4), 648-687.
- Parchomovsky, G. (2000). "Publish or perish", *Michigan Law Review*, 98, 926-951.
- Penin J. (2007). "Open knowledge disclosure: An overview of the evidence and economic motivations", *Journal of Economic Surveys*, 21, 326-348.
- Reinganum J.F. (1982). "A dynamic game of R&D: patent protection and competitive behavior", *Econometrica*, 50, 671-688.
- Reitzig, M. (2004). "Improving patent valuations for management purposes - validating new indicators by analyzing application rationales", *Research Policy*, 33, 939-957.
- Rinner, T.E. (2003). "Protecting minor improvements on core patents: complementing traditional patent protection with strategic disclosure", *The John Marshall Review of Intellectual Property Law*, 2, 398-422.
- Sampat, B.N., D. Mowery and A. Ziedonis (2003). "Changes in university patent quality after Bayh-Dole: A re-examination", *International Journal of Industrial Organization*, 21, 1371-1390.
- Science Watch (2001). "World institutional rankings in select fields, by total citations, 1991-2001", July/August 2001, http://archive.sciencewatch.com/july-aug2001/sw_july-aug2001_page2.htm
- Schmoch, U. (2008). "Concept of a technology classification for country comparison." Final report to the World Intellectual Property Organization (WIPO), Fraunhofer Institute for Systems and Innovation Research, Karlsruhe.
- Silverman A.B. (2003). "Duty to disclose prior art to the United States Patents and Trademarks Office", *JOM Journal of the Minerals Metals and Materials Society*, 55(10), 64-64.
- Spencer, J.W. (2003). "Firm's knowledge-sharing strategies in the global innovation system: empirical evidence from the flat panel display industry", *Strategic Management Journal*, 24(3), 217-233.
- Stokes, D.E. (1997). "Pasteur's quadrant: Basic science and technological innovation", Brookings Institute, Washington D.C.

- Trajtenberg M, A. Jaffe and R. Henderson (1997). "University versus corporate patents: A window on the basicness of invention", *Economics of Innovation and New Technology*, 5, 19-50.
- Van Vianen B., H. Moed and A. Van Raan (1990). "An exploration of the science base of recent technology", *Research Policy*, 19, 61-81.
- Webb, C., D. Harhoff, and K. Hoisl (2005). "Analyzing European and international patent citations: a set of EPO database building blocks". OECD Science, Technology and Industry Working Paper no. 2005/9, Paris
- Winter, S. (2006). "The logic of appropriability: from Schumpeter to Arrow to Teece." *Research Policy*, 35, 1100-1106.
- Ziedonis R. H. (2004). "Don't fence me in: fragmented markets for technology and the patent acquisition strategies of firms." *Management Science*, 50, 804-820

Appendix

Table 7 about here

Table 8 about here

Table 9 about here

Tables

Table 1: Descriptive statistics

	all patent applications		Granted		rejected		t-test	
	mean	std. dev.	mean	std. dev.	Mean	std. dev.	mean diff.	
REF	4.44	2.98	4.51	2.92	4.34	3.05	0.17	**
Log(REF)	1.60	0.60	1.63	0.54	1.56	0.66	0.07	***
XY REF	1.39	1.91	1.22	1.78	1.61	2.04	-0.40	***
XY REF/REF	0.24	0.30	0.22	0.28	0.27	0.32	-0.05	***
NPL	0.46	1.46	0.43	1.47	0.51	1.45	-0.08	**
JOUR	0.26	0.86	0.23	0.82	0.28	0.83	-0.05	**
XY JOUR REF / XY REF	0.05	0.19	0.04	0.18	0.06	0.21	-0.02	***
X JOUR REF / XY REF	0.03	0.14	0.02	0.13	0.03	0.15	-0.01	***
Y JOUR REF / XY REF	0.02	0.13	0.02	0.12	0.03	0.14	-0.01	***
XY comp JOUR REF/XY JOUR REF	0.04	0.17	0.03	0.15	0.05	0.20	-0.02	***
X comp JOUR REF/XY JOUR REF	0.04	0.25	0.03	0.21	0.05	0.28	-0.02	***
Y comp JOUR REF/XY JOUR REF	0.02	0.16	0.01	0.13	0.03	0.20	-0.02	***
IPCs	2.53	1.56	2.57	1.56	2.48	1.57	0.09	**
TIME	4.98	2.33	5.26	2.20	4.59	2.43	0.67	***
INV	2.72	1.77	2.78	1.82	2.65	1.68	0.13	***
PUBLIC	0.01	0.12	0.02	0.13	0.01	0.11	0.01	
NPLs before 1988	0.03	0.17	0.03	0.17	0.03	0.17	-0.00	
Unidentified NPLs	0.07	0.25	0.06	0.24	0.07	0.26	-0.01	**
# obs	6,226		3,559		2,667			

Table 2: Heteroscedastic Probit model for the patent grant decision (6,226 obs)

	Model 1	Model 2	Model 3	Model 4	Model 5
	coeff (s.e.)	coeff (s.e.)	coeff (s.e.)	coeff (s.e.)	coeff (s.e.)
Log(REF)	0.37** (0.18)	0.42** (0.19)	0.50** (0.22)	0.52** (0.22)	0.52** (0.22)
XY REF/REF		-1.16*** (0.32)	-1.29*** (0.35)	-1.29*** (0.35)	-1.28*** (0.35)
XY JOUR REF/ REF			-1.41*** (0.52)	-0.91 (0.56)	
X JOUR REF/ REF					-0.94 (0.65)
Y JOUR REF/ REF					-0.77 (0.75)
XY comp JOUR REF/ REF				-0.97** (0.48)	
X comp JOUR REF/ REF					-0.61 (0.56)
Y comp JOUR REF/ REF					-1.49** (0.62)
X REF/REF					
Y REF/REF					
Log(IPCs)	0.46*** (0.14)	0.39*** (0.13)	0.39*** (0.13)	0.40*** (0.13)	0.40*** (0.13)
Log(TIME)	1.92*** (0.51)	1.75*** (0.45)	1.80*** (0.47)	1.85*** (0.49)	1.84*** (0.50)
Log(INV)	0.30* (0.16)	0.26* (0.15)	0.27* (0.16)	0.28* (0.16)	0.28* (0.16)
Public	-0.02 (0.48)	-0.10 (0.46)	-0.05 (0.47)	-0.06 (0.49)	-0.06 (0.48)
constant	-3.02*** (0.99)	-2.65*** (0.89)	-2.71*** (0.92)	-2.87*** (0.98)	-2.85*** (0.99)
LR-tests:					
Appl. Years	78.65***	78.64***	70.07***	74.58***	75.04***
Priority dummies	248.22***	243.59***	244.44***	252.54***	252.46***
Tech. classes	34.33***	36.14***	36.80***	35.35***	34.97***
Wald test for heteroscedasticity	90.83***	84.58***	82.48***	85.38***	85.36***
LogLikelihood	-3,889.26	-3,865.17	-3,851.24	-3,847.42	-3,846.40

***, **, * indicate statistical significance at the 1%-, 5%-, 10%-level.

Standard errors are clustered for the individual applicants.

Estimated coefficients for a dummy for the share of NPLs that could not be identified (model 3, 4 and 5) and a dummy for NPLs issued before 1988 (model 4 and 5).

Table 3: List of top 25 Journals

JOURNALS	TOTAL SAMPLE			CORPORATE			NON-PROFIT		
	Articles	Share	Rank	Articles	Share	Rank	Articles	Share	Rank
APPLIED PHYSICS LETTERS	354	0.166	1	113	0.155	1	200	0.17	1
JAPANESE JOURNAL OF APPLIED PHYSICS	150	0.071	2	79	0.108	2	64	0.054	4
JOURNAL OF APPLIED PHYSICS	132	0.062	3	35	0.048	5	72	0.061	2
IEEE TRANSACTIONS ON ELECTRON DEVICES	103	0.048	4	60	0.082	3	33	0.028	8
THIN SOLID FILMS	95	0.045	5	21	0.029	11	65	0.055	3
IEEE ELECTRON DEVICE LETTERS	80	0.038	6	32	0.044	6	33	0.028	8
JOURNAL OF THE ELECTROCHEMICAL SOCIETY	75	0.035	7	38	0.052	4	30	0.025	11
JOURNAL OF VACUUM SCIENCE & TECHNOLOGY A-VACUUM SURFACES AND FILMS	74	0.035	8	28	0.038	7	35	0.03	7
JOURNAL OF CRYSTAL GROWTH	73	0.034	9	22	0.030	10	42	0.036	5
ELECTRONICS LETTERS	68	0.032	10	23	0.032	9	38	0.032	6
MATERIALS SCIENCE & ENGINEERING	40	0.019	11	5	0.007	23	33	0.028	8
IEEE PHOTONICS TECHNOLOGY LETTERS	37	0.017	12	15	0.021	12	19	0.016	15
SOLID STATE TECHNOLOGY	34	0.016	13	26	0.036	8	5	0.004	33
MICROELECTRONIC ENGINEERING	34	0.016	13	15	0.021	12	17	0.014	16
SOLAR ENERGY MATERIALS AND SOLAR CELLS	28	0.013	15	7	0.001	17	21	0.018	12
SOLID-STATE ELECTRONICS	26	0.012	16	5	0.007	23	20	0.017	14
SENSORS AND ACTUATORS A-PHYSICAL	24	0.011	17	5	0.007	23	17	0.014	16
SEMICONDUCTOR SCIENCE AND TECHNOLOGY	23	0.011	18	7	0.001	17	14	0.012	19
PHYSICA C	23	0.011	18	2	0.003	38	21	0.018	12
NATURE	22	0.010	20	6	0.008	19	11	0.009	22
SCIENCE	21	0.001	21	5	0.007	23	13	0.011	20
ADVANCED MATERIALS	18	0.009	22	2	0.003	38	16	0.014	18
IEICE TRANSACTIONS ON ELECTRONICS	17	0.008	23	15	0.021	12	2	0.002	66
IEEE JOURNAL OF SOLID-STATE CIRCUITS	17	0.008	23	10	0.014	16	6	0.005	31
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SECTION	17	0.008	23	6	0.008	19	11	0.0093	22

Table 4: Tests for different average citation lags (in years)

Results of the t-Test with unequal sample size and unequal variance				
	Observations	Mean	Std. Err.	t-Test
Public	1413	4.120	0.067	
Companies	962	4.674	0.09	
Mean difference		-0.553	0.110	t = -5.013

N.B.: We consider only articles which are exclusively authored by either corporate scientists or from public institutes. No co-authored articles are included. All citations to patents are included.

Table 5: Tests for different distributions of reference categories

Level of Aggregation	Categories	Companies	Public	Total
Aggregated	Others	463	683	1146
	X or Y	499	730	1229
	Total	962	1413	2375
	Pearson- X^2 (1) = 0.0099			Pr = 0.921
Disaggregated	Others	464	683	1147
	Xs	324	515	839
	Ys	174	215	389
	Total	962	1413	2375
	Pearson- X^2 (2) = 4.1233			Pr = 0.127

N.B.: We consider only articles which are exclusively authored by either corporate scientists or from public institutes. No co-authored articles are included. All citations to patents are included.

Table 6: Citations of public and corporate publications in science (4,982 obs)

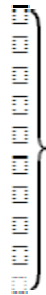
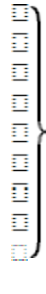
	Poisson with robust standard errors		Negative Binomial	
	coeff. (s.e.)	coeff. (s.e.)	coeff. (s.e.)	coeff. (s.e.)
Public	-0.05 (-0.07)	-0.06 (-0.07)	-0.06 (-0.04)	-0.07 (-0.04)
Cited in Patents	0.64 *** (-0.18)	0.86 ** (-0.33)	0.51 *** (-0.13)	0.32 * (-0.18)
Public * Cited in Patents		-0.15 (-0.36)		0.4 (-0.27)
Page number	0.11 *** (-0.01)	0.11 *** (-0.01)	0.16 *** (-0.01)	0.16 *** (-0.01)
LR-tests:				
Pub. Years	1751.95 ***	1753.98 ***	36.13 ***	35.35 ***
Journals	28779.46 ***	28795.37 ***	1071.60 ***	1073.56 ***
Loglikelihood	-75,723.5	-75,715.5	-18,138.1	-18,137.1

***, **, * indicate statistical significance at the 1%, 5% and 10%-level.

Standard errors are clustered for journal issues.

All regressions include 12 publication year dummies and 32 journal dummies.

Table 7: Distribution of our sample of patent applications across different technology classes within semiconductor

6-digit IPC class	Description	# patent assignments	subfield dummies
H01L21	Processes or apparatus specially adapted for the manufacture or treatment of semiconductor or solid state devices	4.412	t1
H01L23	Details of semiconductor or other solid state devices	1.729	t2
H01L25	Assemblies consisting of a plurality of individual semiconductor or other solid state devices	263	t3
H01L27	Devices consisting of a plurality of semiconductor or other solid-state components formed in or on a common substrate	1.422	t4
H01L29	Semiconductor devices specially adapted for rectifying. amplifying. oscillating or switching and having at least one potential-jump barrier or surface barrier; capacitors or resistors with at least one potential-jump barrier or surface barrier	1.129	t5
H01L31	Semiconductor devices sensitive to infra-red radiation. light. electromagnetic radiation of shorter wavelength. or corpuscular radiation and specially adapted either for the conversion of the energy of such radiation into electrical energy or for the control of electrical energy by such radiation; processes or apparatus specially adapted for the manufacture or treatment thereof	666	
H01L33	Semiconductor devices with at least one potential-jump barrier or surface barrier specially adapted for light emission; processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	373	
H01L35	Thermoelectric devices comprising a junction of dissimilar materials; processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	72	
H01L37	Thermoelectric devices without a junction of dissimilar materials; thermomagnetic devices; processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	17	
H01L39	Devices using superconductivity or hyperconductivity; processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	176	t6
H01L41	Piezo-electric elements in general; electrostrictive elements in general; magnetostrictive elements in general; processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	406	
H01L43	Devices using galvano-magnetic or similar magnetic effects; processes or apparatus specially adapted for the manufacture or treatment thereof	55	
H01L45	Solid state devices specially adapted for rectifying amplifying oscillating or switching without a potential-jump barrier or surface barrier; Ovshinsky-effect devices; processes or apparatus specially adapted for the manufacture or treatment thereof	19	
H01L47	Bulk negative resistance effect devices; processes or apparatus specially adapted for the manufacture or treatment thereof	4	
H01L49	Solid state devices not provided for in groups H01L 27/00-H01L 47/00 and H01L 51/00 and not provided for in any other subclass; Processes or apparatus specially adapted for the manufacture or treatment thereof or of parts thereof	8	t7
H01L51	Solid state devices using organic materials as the active part. or using a combination of organic materials with other materials as the active part; Processes or apparatus specially adapted for the manufacture or treatment of such devices	174	t8
Total		10.925	

The total number of patent assignments corresponds to our final sample of 6.380 patent applications. Patent can be assigned to more than one technology class within semiconductor.

Table 8: Bivariate Correlations

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	Granted	1.00											
2.	Log(REF)	0.06	1.00										
3.	XY REF/REF	-0.10	0.16	1.00									
4.	XY JOUR REF/XY REF	-0.05	0.11	-0.11	1.00								
5.	X JOUR REF/XY REF	-0.04	0.07	-0.07	0.76	1.00							
6.	Y JOUR REF/XY REF	-0.03	0.08	-0.10	0.69	0.06	1.00						
7.	XY comp JOUR REF/XY JOUR REF	-0.06	0.12	-0.01	0.63	0.50	0.42	1.00					
8.	X comp JOUR REF XY JOUR REF	-0.04	0.10	0.00	0.49	0.61	0.08	0.77	1.00				
9.	Y comp JOUR REF/ XY JOUR REF	-0.05	0.09	-0.04	0.46	0.07	0.63	0.62	0.15	1.00			
10.	Log(IPCs)	0.04	0.20	-0.01	0.04	0.03	0.03	0.04	0.03	0.02	1.00		
11.	Log(TIME)	0.17	0.01	-0.03	0.06	0.05	0.03	0.03	0.03	0.02	0.03	1.00	
12.	Log(INV)	0.04	0.04	-0.01	0.04	0.05	0.00	0.03	0.04	0.01	0.07	0.04	1.00
13.	PUBLIC	0.01	0.03	-0.03	0.04	0.03	0.03	0.01	0.02	-0.01	-0.02	-0.01	0.06

The bivariate correlations for the main regressors are very low. High values appear only for variables that are correlated by definition, like XY JOUR REF/ XY REF and X JOUR REF/ XY REF or Y JOUR REF/ XY REF, where the former variable is the sum of the latter two. Note that these variables are not used in the same regression model.

Table 9: Heteroscedastic Probit model for the patent grant decision for the subsample of firms with less than 200 patent applications (5,095 obs)

	Model 1	Model 2	Model 3	Model 4	Model 5
	coeff (s.e.)	coeff (s.e.)	coeff (s.e.)	coeff (s.e.)	coeff (s.e.)
Log(REF)	0.28 (0.20)	0.31 (0.20)	0.37* (0.22)	0.37* (0.22)	0.37* (0.22)
XY REF/REF		-0.85*** (0.27)	-0.94*** (0.28)	-0.94*** (0.28)	-0.93*** (0.28)
XY JOUR REF/ REF			-1.05** (0.48)	-0.48 (0.50)	
X JOUR REF/ REF					-0.74 (0.53)
Y JOUR REF/ REF					-1.65** (0.76)
XY comp JOUR REF/ REF				-1.18** (0.55)	
X comp JOUR REF/ REF					-0.75 (0.63)
Y comp JOUR REF/ REF					-0.18 (0.63)
Log(IPCs)	0.43*** (0.15)	0.37*** (0.14)	0.38*** (0.14)	0.39*** (0.15)	0.39*** (0.15)
Log(TIME)	2.01*** (0.51)	1.87*** (0.47)	1.93*** (0.48)	1.99*** (0.50)	1.95*** (0.49)
Log(INV)	0.24* (0.14)	0.21 (0.13)	0.22 (0.14)	0.22 (0.14)	0.22 (0.14)
Public	-0.07 (0.47)	-0.11 (0.46)	-0.08 (0.47)	-0.11 (0.50)	-0.11 (0.49)
Constant	-2.74*** (0.97)	-2.46*** (0.90)	-2.52*** (0.91)	-2.70*** (0.97)	-2.63*** (0.95)
LR-tests:					
Appl. Years	37.54***	39.62***	34.22***	33.54***	24.92***
Priority dummies	120.50***	116.75***	117.99***	116.68***	107.15***
Tech. classes	28.62***	29.59***	28.67***	28.32***	19.38***
Wald test for heteroscedasticity	78.57***	72.89***	71.12***	71.65***	71.16***
LogLikelihood	-3,188.83	-3,176.99	-3,169.98	-3,165.88	-3,165.22

***, **, * indicate statistical significance at the 1%-, 5%-, 10%-level.

Standard error are clustered for the individual applicants.

Estimated coefficients for a dummy for the share of NPLs that could not be identified (model 3, 4 and 5) and a dummy for NPLs issued before 1988 (model 4 and 5).

All regressions include five application year dummies, eight technology class dummies and four priority dummies.