Why Weak Patents?

Rational Ignorance or Pro-“customer” Tilt?

Zhen Lei and Brain D. Wright

Department of Agricultural and Resource Economics

UC Berkeley

207 Giannini Hall, Berkeley, CA 97720

zlei@berkeley.edu and wright@are.berkeley.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29, 2009

Why Weak Patents?

Rational Ignorance or Pro-"Customer" Tilt?

Zhen Lei and Brian Wright

Abstract

The issuance of weak patents is widely viewed as a fundamental problem in the current US patent system. Reasons that have been offered for the granting of weak patents by the US Patent and Trademark Office (USPTO) include examiners' "rational ignorance" of the patentability of applications and pro-"customer" rules and institutions that create incentives for examiners to grant patents of dubious validity to their "customers"- applicants. In this paper, we study whether US examiners' behavior in prior art search betrays their assessment of applications' patentability. For a sample of US patents for which applications were also filed at the European Patent Office (EPO), we construct a measure of the fraction of prior art that is missed by US examiners. We find that this measure significantly explains the probability of receiving a patent at the EPO. The results are robust to different empirical specifications. US examiners' prior art searches indicate that they are, on average, not "rationally ignorant". On the contrary, they identify and dedicate more search effort to those applications that seem more problematic, because they bear the burden of proof of non-patentability. Our study offers empirical evidence that a systematic problem of weak patents likely exists, and suggests that the problem may be more strongly attributable to the pro-applicant rules and policies than to examiners' ignorance. The current prevalence of weak patents does not appear to be caused at the margin by lack of resources at the USPTO.

---

* Department of Agricultural and Resource Economics at UC Berkeley, zlei@berkeley.edu and wright@are.berkeley.edu.
This paper has benefited from discussions with Maximilian Auffhammer, Michael Anderson, Robert Barr, Peter Berck, Sara Boettiger, Gregory Graff, Stuart Graham, Dietmar Harhoff, Ann Harrison, Ethan Kaplan, Larry Karp, Alexandre Mas, David Mowery, Jeffrey Perloff, Elisabeth Sadoulet, Guanming Shi, Ted Sichelman, Sofia Villas-Boas, Michael Ward, Catherine Wolfram, David Zilberman, and numerous seminar participants. Paul Jensen, Alfons Palangkaraya and Elizabeth Webster generously provided their dataset. M-CAM, Inc. kindly offered their linguistic and semantic mapping service. I also thank Kyle Jensen for collecting some patent information.

1. Introduction

The US patent system has a key role in stimulating technological innovation by offering the prospect of economic rewards to inventors, and patenting is a big business in the United States. In 2007 approximately 467,243 patent applications were filed with the US Patent and Trademark Office (USPTO) and 184,377 new patents were issued, and both numbers have been growing steadily.[1]

However, there has been increasing concern about the operation of the US patent system in recent years. Critics such as Adam Jaffe and Josh Lerner claim that "the system is broken and endangering innovation and progress," and one of the fundamental problems is that "the technology world today is awash in patents that should not have been granted in the first place."[2] Patents " that should not have been granted in the first place" are called weak patents, because they are either not novel or obvious, in the light of *prior art* (including prior inventions and knowledge).[3] Another way to think about weak patents is the notion of patent weakness (strength), conceptualized as the probability of being invalidated (validated) under a perfect re-examination[4] or by an ideal court trial[5] if challenged.

What are the costs of weak patents?[6] Since market power enjoyed by a patent holder is typically considered to be a social cost that is necessary to stimulate innovation, weak patents might cause social costs without commensurate social benefits associated with increased innovation. Furthermore, there is generally no reason to expect that private incentives to challenge weak patents through litigation line up well with the social incentives.[7]

---

1 USPTO Annual Report (2007), Tables 2 and 6 (http://www.uspto.gov/web/offices/com/annual/2007).
2 Jaffe and Lerner (2004), page
3 Some widely cited and absurd patents include a method for swinging on a swing (US Patent No. 6,368,227); and a "sealed crustless sandwich" (US Patent No. 6,004,597).
4 By a perfect re-examination, we mean that the re-examination is a perfectly thorough and accurate implementation of the relevant statutes.
5 It is well recognized that, in practice, a court's decision on patent validity is to a large extent determined by which side has more resources and means to invest on the trial, such as hiring better lawyers, etc. By an ideal court trial we mean that the court decision is immune from those factors (say, in the case of both sides having infinite resource and money), and completely based on the patentability of a patent.
6 For detailed discussions of costs of weak patents, see Lemley and Shapiro (2005) and Shapiro (2004). To give an example, Research In Motion, the company that sells BlackBerry wireless e-mail service, was sued by a company named NTP for infringing NTP's patents and paid a $612 million "ransom" to settle the case, even though the USPTO had already indicated that it was likely to conclude soon that the NTP patents BlackBerry was accused of infringing were not even valid.
7 See Farrell and Merges (2004) for a discussion of various incentives to challenge patents. They point out that private incentives could be suboptimal due to the problems of free-riding and pass-through of royalties.

Proliferation of weak patents in a technology field raises the cost of licensing, litigation, or avoiding infringement.[8] Uncertainty about the validity of previously issued patents may deter investment in innovation and/or distort its direction. Rent-seeking entrepreneurs may divert resources from productive activities into speculative patent acquisition and enforcement ventures. The Patent Reform Act, which has been discussed on the Capitol Hill since 2005, includes measures addressing the problem of weak patents.[9]

Why are weak patents issued? Statutory requirements for patentability of an invention include *novelty, non-obviousness and reduction to practice*.[10] Upon receiving an application, a USPTO examiner is responsible for searching for and obtaining its prior art, and comparing the application to the prior art to evaluate its patentability. Two reasons, broadly speaking, have been offered for why US examiners issue so many weak patents.[11] The first is examiners' lack of qualifications, experience,[12] time and resources for searching for prior art and evaluating each application,[13] all of which are related to the level of funding available to the USPTO.[14] It is alleged that, at the USPTO, an inexperienced workforce of examiners

---

Choi (2005) discusses the discrepancy in private incentives and social incentives to litigate, because of the possibility of cross-licensing and patent pools.

[8] Farrell and Shapiro (2008) study how the structure and level of royalties depends on a patent's strength. They show that when downstream licensees do not compete against each other or the patent holder, license fees approximate the license fee for an ironclad patent times the patent strength. When downstream users compete, they will accept surprisingly large per-unit royalties.

[9] The House passed a version of the patent reform bill (HR 1908) in July, 2007. The bill suffered a serious. setback in May 2008 when its Senate version (S 1145) was stalled. See more of the history of this Patent Reform Act, visit http://www.fr.com/news/articledetail.cfm?articleid=490.

[10] The US Patent Act of 1952 (35 U.S.C.) allows the granting of patents to "any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement" (Section 101), and precludes patent granting for subject matter that "was known or used by others" (Section 102), or "would have been obvious at the time the invention was made to a person having ordinary skill in the art." (Section 103) These are called the utility, the novelty and the non-obviousness requirements, respectively.

[11] See National Academies of Science (hereafter NAS) Study (2004) and Jaffe and Lerner (2004) for comprehensive discussions about why the USPTO might issue so many weak patents. The other side of the equation is that applicants submit too many applications and many of them are bad. The Court of Appeals for the Federal Circuit (CAFC) significantly broadened and strengthened the rights of patent holders. The rate of patent application filings in the US has been accelerating (Jaffe and Lerner, 2004).

[12] It is argued that USPTO cannot retain good examiners and that most examiners have only 2-3 years of examination experience. Also examiners might be unfamiliar with new technologies and lack the knowledge of where to look for prior art.

[13] The average time allocated for an examiner to address one application is understood to be between sixteen and seventeen hours. The NAS study (2004, p51) reports that the number of examiners has not kept pace with the number and complexity of patent applications. The number of examiners per 1000 applications has been declining, while applications have become more complex, as measured by the number of claims and prior art citations per application.

[14] See Federal Trade Commission (hereafter FTC) Report (2003).

conducts superficial searches and overlooks prior art that could render weak patents unpatentable. This situation is not unanimously viewed as problematic. One influential legal scholar has argued that US examiners "are 'rationally ignorant' of the objective validity of patents,…, because it is too costly for them to discover those facts" (Lemley, 2001). He argues that, given the skewed nature of patent value, society is better off if economizing on USPTO examinations and deferring rigorous determination of validity until the patent enters litigation.[15] An alternative reason for the issuance of weak patents is an alleged *pro-applicant* bias of policies and procedures at the USPTO. Critics observe that since the 1980's the USPTO's culture, mission and incentives have been re-oriented towards issuing patents and serving the interests of patent applicants (denoted "*customers*" by the USPTO), and that US examiners are encouraged by various institutional incentives and protocols to accept applications that they perceive to be ineligible (Jaffe and Lerner, 2004).

In this paper, we investigate whether US patent examiners are "rationally ignorant", or whether they can and do distinguish strong applications from weak ones and conduct their examinations accordingly. Specifically, we ask whether a US examiner's behavior in prior art search betrays his assessment of an application's patentability. The basic idea is the following: if an examiner initiating two independent applications is "rationally ignorant," then it is plausible that he spends the same, and limited, amount of time on each application and stops his search when the search time runs out, regardless of their patentability. Understandably, in neither of the two cases is his search thorough; and due to the stochastic nature of the results of prior art searches,[16] one resulting patent likely has a higher share of prior art *missed* by the examiner than the other. We would then expect the patent with a *higher* share of missed prior art to be *weaker*, in the sense that it is more likely that prior art exists and be discovered later and used to invalidate it. Indeed, a view commonly held by scholars and practitioners is that if a patent has a *larger* amount of cited prior art, it has a *higher* likelihood of being validated.[17]

---

[15] Lemley (2001) argues that strengthening the examination process is not cost effective, since very few patents are actually litigated or licensed; most patents simply sit on a shelf unused, or are used only for noncontroversial purposes, like financing. Because of this, society would be better off spending its resources on more judicial inquiries of validity of those few cases in which it matters, rather than paying for more protracted examination of all patents ex ante.

[16] For instance, a prior art search might miss a larger share of prior art if it is more difficult to find the application's prior art or it happens that the examiner conducts the search on Friday afternoon.

[17] Moore (2003) argues that "patents that include more citations or more diverse citations are more likely

Suppose instead that the examiner, constrained though he may be, can actually differentiate between the relative merits of the two applications. How would he decide his search effort? The examiner bears the burden of proof of non-patentability and the proof is the prior art he searches and obtains.[18] If he is going to grant an application, *no proof of patentability* is needed; but if to reject an application, he needs to provide evidence of prior art to show the application is unpatentable. Furthermore, under the various legal, institutional and cultural incentives and constraints at the USPTO, which we shall discuss in Section 2, the examiner might target the application he deems as less patentable. He might give an easier pass to the good application, spending less time on prior art search, so that the resulting patent might have a *higher* share of missed prior art. But for the application he considers weak, the examiner, pursuing a rejection, searches harder for prior art to provide *proof of non-patentability* and to show the applicant the difficulty she would face if she chose to persist in pursuing the patent. In the end, the weak application is eventually allowed because the applicant is persistent and the examiner finally concedes. And we would observe a *smaller* share of missed prior art associated with the resulting patent.

Therefore, if USPTO examiners are "rationally ignorant," a patent with a *higher* share of missed prior art might be *weaker*. However, if examiners do distinguish the relative patentability of applications and allocate their search effort accordingly, a *higher* share of missed prior art would suggest a *stronger* patent.

How do we test the relationship between the share of missed prior art and the strength of a patent? We study a sample of US patents with a USPTO filing date between 1990 and 1995, for which applications were also filed in the Europe Patent Office (EPO). Outcomes from independent EPO application process are used as indirect indicators of these US patents' strength (patentability).

For each US patent *i* in the sample, a Natural Language Processing (NLP) algorithm, which involves linguistic and semantic analysis, is used to retrieve prior patents that are linguistically and likely technically similar to, *but not cited by,* the root patent *i*. We define

---

to be valid."

[18] USPTO practice requires that examiners articulate their reasons for a rejection, however, examiners often say nothing if they chose to allow a case. It is argued that this practice encourages examiners to allow rather than to reject applications.

these uncited prior patents as a measure of patent *i*'s missed prior patents. Based on the information on the number of missed prior patents (MPP_i) and cited prior patents (CPP_i), we construct a metric, *share of missed prior patents* ($SMPP_i$), where $SMPP_i = MPP_i/(MPP_i+CPP_i)$, to proxy for its US examiner's effort in searching for prior art.

We find that, for the US patents in our sample, the share of missed prior patents (SMPP) can significantly explain the probability of non-withdrawal by applicants and the probability of being granted (conditional on non-withdrawal) at the EPO, suggesting that a US patent with a *higher* SMPP is indeed a *stronger* patent. The most convincing result is from a panel data model with *US examiner by technology by US application year* fixed effects. It allows us to identify the effect of SMPP through variations in SMPPs within the set of patents examined by the same US examiner, in the same technology field and with the same USPTO application year. Similar results also hold when we use a panel data model with *US examiner by patent assignee* fixed effects to control for the possible heterogeneous interactions between applicants and US examiners. We also test several alternative rationales for the finding that a higher SMPP is positively correlated with the likelihood of an EPO success, and find that they are unpersuasive.

Our results indicate that, by and large, US examiners are not "rationally ignorant." Instead, they can and do distinguish strong applications from weak ones, and conduct their searches accordingly. This appears consistent with the postulation in Allison et al. (2004) that "the much-maligned PTO is doing a better job than expected in evaluating the patents that really matters," after they find that litigated patents cite more prior art and have a longer USPTO examination than a random sample of general patents.

Our study offers a piece of empirical evidence that the problem of weak patents in the US might be broad and systematic rather than anecdotal. For almost 35% of US patents in our sample, their related applications at the EPO failed, and many of these failures are predicted by the actions of examiners at the USPTO, who recognized their lack of patentability and attempted to reject them by putting more effort in search for prior art. This suggests that the USPTO indeed issues patents weaker, on average, than its European counterpart.

Our study suggests that, in regard to reasons for issuing weak patents, the pro-"customer" policy and institutions, which make it difficult for US examiners to reject ineligible applications, might be more salient than examiners' ignorance. Even without changing the number of US examiners and the workload and the time allocated for each application, the weak patent problem might be significantly addressed just by empowering examiners to be able to reject applications that they consider as non-patentable. A tradeoff between weeding out weak patents and increasing USPTO expenditure, suggested by the "rational ignorance" theory in Lemley (2001), does not appear to exist on the margin.

The remainder of the paper is organized as follows. Section 2 offers some institutional details at the USPTO to develop our hypothesis about examiner's search effort and its implication for patent strength. Section 3 provides a simple model to illustrate the hypothesis. In Section 4 we describe the data and our empirical strategies for testing our hypothesis. We present the results in Sections 5 and propose a novel metric for patent strength in Section 6. Section 7 discusses the policy implications of our findings and concludes.

2. Examiners, Examination Procedures and Incentives at the USPTO

In this section, we briefly describe some institutional details about examiners, the patent examination process, and institutions and incentives for examiners at the USPTO, to develop our hypothesis that if US examiners are not "rationally ignorant" and distinguish stronger applications from weak ones, they put more effort in search for prior art for applications that they think are less patentable. As a result, a patent with a *smaller* share of missed prior art might be weaker.

2.1 Examiners and their burden of proof at the USPTO[19]

The USPTO is currently staffed by over 5477 patent examiners, and has more than 8900 total full-time equivalent employees. Examiners work together on closely related subjects in small groups called *art units*. Several related art units are organized into a work group, and several work groups covering a wide technology area are grouped into a

[19] See Cockburn et al. (2003) for more detailed and excellent institutional discussions of the USPTO.

technology center. The USPTO has eight technology centers[20] and approximately 271 art units.

Art units are the building blocks of the US patent examination system. Each art unit is led by a Supervisory Patent Examiner and contains 10–15 primary and assistant examiners. Primary examiners have at least 5 years of experience at the USPTO and have signatory authority in granting or rejecting patents. Assistant examiners are junior examiners who are like apprentices and must have their examinations reviewed and signed by a primary examiner. It takes five to six years for assistant examiners to become primaries.

The workflow for the patent application process is quite systematic. After being received at a central receiving office and passing basic checks to qualify for a filing date, patent applications are sorted by the Office of Initial Patent Examination, which allocates them to one of the art units. Within the art unit, the Supervisory Patent Examiner looks at the invention claimed in the application and assigns it to a specific examiner. The assigned examiner will, in most cases, have continuing responsibility for the examination of the application. He reads and understands the application, and searches for prior art (including previous patents, databases, and journals) to determine if this is patentable, i.e., if it is a novel, non-obvious and useful invention.[21] A prior art search typically begins with a review of existing US patents in relevant technology classes and subclasses, either through computerized tools or by manual examination of hard-copy stacks of issued patents, and may then proceed to a search of foreign patent documents, scientific and technical journals, or other databases.

Is the intensity of search independent of the examiner's judgment of the application? An answer is suggested by the following quote from a patent attorney representing

[20] The eight technology centers at the USPTO are: 1600 Biotechnology, Organic Chemistry; 1700 Chemical and Materials Engineering; 2100 Computer Architecture, Software, and Electronic Commerce; 2600 Communications; 2800 Semiconductors, Electrical and Optical Systems and Components; 3600 Transportation, Construction, Agriculture, National Security and License and Review; 3700 Mechanical Engineering, Manufacturing, and Products; and 2900 Designs for Articles of Manufacture.

[21] When an application is filed, the applicant has *no* duty to do a thorough search for prior art and disclose them to the USPTO, although he has a duty to disclose the prior art of which he is aware that is "material to the patentability of the application." It is the examiner's responsibility to search for prior art to determine whether the claimed invention is patentable.

applicants,[22]

"I do not have to prove my invention is patentable. It is the examiner who has to prove my invention is unpatentable."

As suggested, an examiner's responsibility is to prevent unpatentable applications from being granted. It is this function that distinguishes a patent examination system from a patent registration system.

Moreover, the examiner is not allowed to use his "gut feeling" to determine an application's patentability, as suggested by the following quote from an ex-examiner:[23]

"I felt very sad when I had a gut feeling about a (bad) application but could not find the prior art (to reject it)."

Two messages can be read from the quotes above. First, examiners do form an opinion about an application's patentability during the examination. Second, when an examiner tries to reject an application, he bears the burden of proof of non-patentability and must articulate his reasons and provide prior art to justify his rejection. By contrast, an allowance of an application does not require him to prove anything. With such burden of proof of non-patentability, an examiner may search more diligently for prior art when he regards an application as unpatentable and considers a rejection.

The Manual of Patent Examining Procedure (MPEP), the official guideline for patent examiners, instructs examiners to search for prior art that not only shows the unpatentability of original claims in an application but also, in a FOAM of non-final rejection, anticipates how the applicant will amend claims in response to the FOAM. The following excerpt is from MPEP Section 904.03 on "Conducting the Search":

---

[22] The quote is from an instructor in a course on patent prosecution that the author audited in the Boalt School of Law at UC Berkeley. The instructor is a patent lawyer who taught law students on how to prosecute patent applications for their clients.

[23] The quote is from an informal conversation with an ex-examiner at the USPTO, who prefers to remain anonymous.

"It is normally not enough that references be selected to meet only the terms of the claims alone, …., the search should, insofar as possible, also cover all subject matter which the examiner reasonably *anticipates* might be incorporated into applicant's amendment." (Italics added)

The MPEP guidelines suggest that for a very good application, an examiner may not need to conduct as thorough a search, because he is about to approve the application, and the applicant is less likely to come back. However, if the examiner thinks an application to be unpatentable and considers a rejection, he needs to search harder for prior art to demonstrate the non-patentability of the original claims and the non-patentability of the anticipated claim amendments. All the burden of proof of non-patentability is on the examiner.

2.2 Incentives and constraints in patent examination at USPTO

After the examiner obtains and reads the prior art, he determines whether the application is patentable, and writes a letter of "*first office action on merit* (FOAM)" to the applicant (or normally, the applicant's attorney) with either a Notice of Allowance or, more commonly, a Non-final Rejection. When writing a FOAM of non-final rejection, the examiner must write a detailed analysis of the basis for the rejection.[24] The applicant then has a fixed length of time to respond by supplying additional arguments and evidence and/or amending the claims. After negotiation, the examiner writes a letter of "*second and final office action*" to allow the application or maintain some or all of the initial rejections. This second rejection is typically called a "final" rejection.

An overwhelming majority of applications at the USPTO receive a FOAM of non-final rejection, but only a minority of them receives a "final" rejection in the second office action. Moreover, a majority of the applications that receive a second and "final" rejection ultimately result in a patent, even without amendment.[25] This seemingly puzzling

[24] The Manual of Patent Examination Procedures (MPEP), Section 706.
[25] Lemley and Sampat (2007) find that in their sample, 86.5% of the applications received a first office action on merit (FOAM) of a non-final rejection, but only 34.5% received a "final" rejection in the subsequent action. Subsequently, 52.9% of those applications that received a "final" rejection ultimately

fact could be, at least partly, explained by the cultural, procedural and institutional incentives at the USPTO that encourage examiners to grant, rather than continue in trying to reject applications from persistent applicants.

The compensation and promotion schemes for individual USPTO examiners provide incentives for them to process applications as quickly as possible, by allowing them. For each examiner, the USPTO sets a productivity goal specifying the number of hours he is to spend on an average application. In practice, the productivity goal is specified as a certain number of points that the examiner is supposed to earn, calculated on a biweekly basis (the "biweekly production goal"). The examiner is awarded a point when he either writes a FOAM or disposes of an application (the application being either abandoned, allowed, or appealed to the Board of Patent Appeals and Interference in which case the examiner needs to write an Examiner's Answer in response). But no points are awarded for all other actions including: (1) a second, third, etc., action on the merits; (2) a final rejection; (3) an interview (in person or by telephone); and (4) an Advisory Action. The biweekly production goal is based on the examiner's technology area and his experience level (primary or assistant examiner). If he exceeds 110% of his production goal, the examiner receives a bonus.[26]

The USPTO also implements various internal assessments to ensure "examination quality control" through auditing an examiner's work. The primary quality indicator is the examiner's error rate.[27] The USPTO quality review specialists calculate this rate by analyzing a sample of allowed patents for patentability issues, such as the adequacy of the examiner's search and the originality of the applicant's claims, and determining the percentage of patents that contain at least one claim that would be held invalid in a court of law.[28] There are informal "controls" of examination quality as well. For instance, the examiner might care

---

result in a patent, and another 20% are still pending. 56.7% of the applications issued are issued without any amendment. 66.1% of those which are amended after a final rejection are patented.

[26] For a detailed discussion of USPTO examiners' biweekly production goal and the reward system see a report (2004) by the Office of Inspector General, US Department of Commerce, (hereafter the OIG Report (2004).

[27] Other indicators include: (1) USPTO reopens applications based upon applications sampled for post review; (2) the Board of Patent Appeals and Interferences reviews adverse patentability decisions; and (3) the examiner's grant rate. (The Office of Inspector General Report (2004), p.14).

[28] The error rate reported in USPTO quality assessment audits rose slightly during the 1990s, but has only ranged between 3.6% and 7% since 1980.

about his reputation and "does not want his patent to be the one that hit newspapers."[29] With these "quality control" mechanisms in place, the examiner likely targets applications that he considers to be unpatentable, for which he searches harder for prior art to prove their unpatentability.

Moreover, certain examination procedures at the USPTO allow an applicant essentially unlimited attempts to persuade a critical examiner to approve a patent, making it impossible for the examiner to ever finally reject a bad application if the applicant is persistent. As Lemley and Moore (2003) observe, "One of the oddest things about the US patent system is that it is impossible for the USPTO to ever finally reject a patent application. While patent examiners can refuse to allow an applicant's claim to ownership of a particular invention, and can even issue what are misleadingly called 'final rejections', the applicant always gets another chance to persuade the patent examiner to change his mind." The term "final rejection" is a classic legal misnomer.[30] The applicant receiving a second and final rejection has several options: 1) she can continue to negotiate with the examiner by submitting claim amendments, evidence and arguments; 2) she can request a face-to-face or telephone interview with the examiner to try to persuade the examiner in person; or 3) she may choose to appeal the rejection. In these cases, the examiner is awarded no points until the application is disposed of. Alternatively, the applicant can start the examination process over by filing a Request for Continuation Examination (RCE), or application continuation.[31]

Furthermore, beginning in the early 1990s, the US Congress converted the USPTO from an agency funded by tax revenue to one funded by the fees it collects. The USPTO has since increasingly viewed itself as an organization whose mission is to serve its "customer", patent applicants who want their applications granted.[32] This new orientation creates strong incentives for the USPTO to process applications as quickly as possible. Examiners have

[29] The quote is from an informal conversation with an ex-examiner at the USPTO.
[30] See Robert P. Merges et al., Intellectual Property in the New Technological Age (p116, 3rd edition, 2003) ("The label 'final rejection' is a misnomer if ever there was one.")
[31] Application continuations permit an applicant to re-file a pending application and avoid the implementation of a patent examiner's decision. There are three types of continuations: Continuation Application, Continuation-In-Part, and Division. Applicants can also file a "Request Continuation Examination". See Lemley and Moore (2003) for a detailed discussion about the problem of application continuation.
[32] See Jaffe and Lerner (2004) for a history of how the USPTO has become more service-oriented.

reportedly been criticized by their supervisors for undertaking too many reviews of prior art before issuing a patent.[33]

With all these institutional and cultural barriers to reject an application, all an examiner can do when attempting to reject an unpatentable application is to search for more prior art to increase the difficulty, delay and cost[34] for an applicant should she choose to persist, and perhaps to persuade her to narrow the claims in the application. The examiner hopes that such deterrence will lead to a more speedy disposal of the application. Should the applicant persist, however, it is costly for the examiner to engage in back-and-forth negotiations with the applicant; it consumes time and does not earn the examiner production points.

Although the examiner might ultimately grant a weak patent due to the applicant's persistence, the fact that he conducted a more thorough search for prior art reflects his view that the application is relatively weak. Thus, the examiner's search effort might be a signal of patent strength. This is the hypothesis explored in this study.[35]

3. A Simple Model

We use a simple model to illustrate the hypothesis presented in the Section 2. Define the underlying patentability of a US application *i* as a random variable P_i, with a normal distribution[36] that is common knowledge:

---

[33] Here are selected and very possibly non-representative quotes from PTO examiners: (1) We have a cultural goal now. If some examiner is not issuing enough, his SPE (supervisory patent examiner) will complain and make her or him feel like s/he's a weirdo anal retentive tight butt…. The examiner wonders why s/he is working so hard. The examiner wonders why s/he draws complaints from the boss…. (2) We just don't fight hard enough against the bull- being shoveled by upper management…. And why should you care? Hey, management pays you for good patents or bad, right? Why should you fight with management? Why reject? (These comments are taken from Gregory Aharonian, "A Few Patent Examiners Complain about Patent Quality," PATNEWS, January 28, 1999.

[34] The lengthy examination is costly for applicants because patent lawyers who represent applicants charge fees according to the hours they spend on the case.

[35] Our hypothesis postulates that if an application is less novel or more obvious, the resulting patent has a smaller share of missed prior art. In another scenario, when an application has excessively broad claims, the examiner might conduct a more thorough search for prior art to narrow down the claims in the issued patent. This scenario could create a bias against our finding that a US patent with a higher share of missed prior art is more likely to succeed at the EPO, as this application with broad claims is likely to be narrowed down and issued at the EPO as well, and thus the share of missed prior art would have no explanatory power in its EPO application outcome.

[36] If we prefer an application's patentability to have a range of $[0, \infty)$, we can consider P_i as the logarithm of the underlying patentability that has a lognormal distribution.

(1) $$f(P_i) \sim N(\alpha, \sigma^2)$$

Suppose that Q_i is the signal about the patentability of the application *i*. Q_i is also a random variable in the sense that people may receive different signals of this patentability. We assume that Q_i has the following conditional distribution (conditional on the true underlying patentability P_i) that is also normal and common knowledge:

(2) $$f(Q_i \mid P_i) \sim N(P_i, \upsilon^2)$$

The US examiner reads the application *i* and receives a signal about its patentability, q_i, which is a realization of the random variable Q_i. We can think of q_i as the examiner's perception about the application's patentability.[37] With this perceived patentability q_i and the common prior belief about the true patentability P_i, the examiner forms his posterior belief about P_i, shown in Equation (3).

(3) $$f(P_i \mid q_i) \sim N(\frac{\sigma^2 q_i}{\sigma^2 + \upsilon^2}, \frac{\sigma^2 \upsilon^2}{\sigma^2 + \upsilon^2})$$

The examiner then decides how much effort he wants to put in search for prior art. The search effort is *e* with a support [0,1].

The USPTO has a patentability threshold p^*. For the examiner, granting an application with a patentability P_i below the threshold p^* incurs a cost, either because the quality assessment office or other people will with some probability detect the mistake later or the examiner might just feel bad for not meeting his job responsibility. We assume a constant cost C_1 for granting an application whose P_i is below p^*.

To try to put an end to an application that the examiner deems unpatentable, he conducts a more diligent search for prior art, not only to show a proof of non-patentability, but perhaps more importantly, to deter the applicant from persisting in pursuing the patent. If the applicant is determined, she can keep coming back over and over again. Suppose that the probability of the applicant being persistent *D* is inversely dependent on the share of prior art the examiner has found, which is in turn determined by the search effort e_i, in a simple form

[37] For simplicity, the model assumes that the examiner receives the signal and decides his search effort once and for all. Alternatively, the examiner, during the search, could do Bayesian updating of his perception of patentability and adjust his search effort accordingly. By contrast, the theory of "rational ignorance" postulates that the examiner does not link his search effort to the patentability of the application.

$D(e_i)=(1-e_i)$.[38, 39]

However, the search effort is costly to the examiner, as a higher search effort consumes more time and thus reduces the points he will earn towards his biweekly production goal. We assume the cost of search effort to be $C(e_i)=C_0\, e_i^2$, where C_0 is a constant. In our study, the examiner's search effort *e* is measured by the share of prior art that is cited in the resulting patent.

Given his perception about application *i*'s patentability q_i, and his posterior belief about the underlying patentability $f(P_i|q_i)$, the examiner decides his search effort, e_i, which minimizes the sum of the cost of search and the expected cost of approving application *i*:

$$(4) \qquad \min \quad C_0 e_i^2 + C_1 D(e_i)\int_{-\infty}^{p^*} f(P_i \mid q_i) d(P_i \mid q_i)$$

where the probability of granting the application *i* when its P_i below p^* is $D(e_i)\int_{-\infty}^{p^*} f(P_i \mid q_i) d(P_i \mid q_i)$, the product of the probability of the applicant being persistent (in which case the examiner has to grant) and the probability of P_i being smaller than p^*.

The examiner solves this cost minimization problem and decides his search effort, e_i, which is given by Equation (5).

$$(5) \qquad e_i = \frac{C_1}{2C_0}\Phi\left(\frac{(\sigma^2+\upsilon^2)p^* - \upsilon^2\alpha - \sigma^2 q_i}{\sigma^2\upsilon^2}\right)$$

where $\Phi()$ is the cdf function for a normal distribution.

The derivative of e_i with respect to q_i is negative, meaning that if the examiner perceives application *i* to have higher patentability, he will use less search effort. Understandably, the examiner will take more search effort if the USPTO sets a higher patentability standard p^*. The examiner's search effort is also higher if the cost of granting an undeserved patent is higher or if the search cost is lower:

$$(6) \qquad \frac{\partial e_i}{\partial q_i} < 0$$

---

[38] Since an applicant can always obtain a patent if he is persistent and willing to spend time and money to contest a rejection, the patentability of the application *per se* does not affect his decision on whether to fight. It is the difficulty and cost of persuading the examiner, reflected by the prior art cited by the examiner, that matter to the applicant.

[39] The probability of the applicant being persistent could also depend on the application's potential commercial value, which we assume away in the model, for simplicity. This simplification does not change the results of the model

(7) $$\frac{\partial e_i}{\partial p^*} > 0, \frac{\partial e_i}{\partial C_1} > 0, \frac{\partial e_i}{\partial C_0} < 0$$

As outside researchers, we observe the prior art cited by the examiner and we have an estimate of missed prior art. These two pieces of information can be used to construct an indicator of the examiner's search effort e_i. The question is, what is the conditional expectation of the true patentability P_i, conditional on e_i (i.e., $E[P_i|e_i]$)?

Given e_i, we can derive q_i, the examiner's perception of the patentability, from Equation (5):

(8) $$q_i(e_i) = \frac{(\sigma^2 + \upsilon^2)p^* - \alpha\upsilon^2}{\sigma^2} - \upsilon^2\Phi^{-1}\left(\frac{2C_0 e_i}{C_1}\right)$$

(9) $$\frac{\partial q_i(e_i)}{\partial e_i} < 0$$

where $\Phi^{-1}()$ is the inverse of the cdf function of a normal distribution.

After we derive $q_i(e_i)$, we can, from the conditional distribution $f(P_i|q_i)$ presented in Equation (3), infer the expectation of the underlying patentability P_i, conditional on e_i, i.e., $E[P_i|e_i]$.

(10) $$E\left[P_i \mid e_i\right] = E\left[P_i \mid q_i(e_i)\right] = \frac{\alpha\upsilon^2 + \sigma^2 q_i(e_i)}{\sigma^2 + \upsilon^2}$$

As shown in Equation (11), the derivative of $E[P_i|e_i]$ with respect to e_i is negative, meaning that if we observe that a US patent has a higher search effort e_i, the conditional expectation of its underlying patentability ($E[P_i|e_i]$) should be lower.

(11) $$\frac{\partial E[P_i \mid e_i]}{\partial e_i} = \frac{\partial E[P_i \mid e_i]}{\partial q_i} * \frac{\partial q_i}{\partial e_i} = \left(\frac{\sigma^2}{\sigma^2 + \nu^2}\right) * \frac{\partial q_i}{\partial e_i} < 0$$

Thus, the model illustrates that if a US examiner has a good assessment about the patentability of an application, he would conduct a less diligent search for prior art if he perceives it to be more patentable. As a result, we would expect a patent with a *higher* share of missed prior art to be a *stronger* patent.[40] On the contrary, if the examiner is "rationally

[40] In this simplified model, we assume that whether an application is granted by the USPTO is only determined by the applicant's persistence, which is in turn determined by the examiner's search effort and

ignorant" and allocates the same amount of time to each application, a patent with a *higher* share of missed prior art might be *weaker*. Therefore, an empirical study of the relationship between a patent's share of missed prior art and its strength would be an indirect test of whether US examiners are "rationally ignorant."

4. Empirical Strategies and Data Description

How do we empirically test whether a higher share of missed prior art, indicates a stronger or weaker patent? First, we need a sample of US patents and some indicators of which are stronger and which are weaker. Second, for each of these US patents, we need to know how much prior art is cited and how much has been missed by its US examiner during the examination, so that the share of missed prior art can be constructed.

4.1 Use of international patenting to distinguish stronger US patents from weaker ones

Ideally, we would like to have a random sample of US patents drawn from the whole US patent pool and have them either examined again with a *perfect* re-examination at the USPTO or litigated for validity in the court with an *ideal* trial. The outcomes of the USPTO re-examinations or the validity decisions by the court would tell us their patent strength.

Since 1981, the USPTO has had a reexamination process available at any time during a patent's life. However, the US reexamination has serious disincentives and drawbacks and is almost dysfunctional.[41] Patent litigation data, which contain only a small fraction of patents that are highly selected, might be too complex and biased to be suitable for our purpose. Only 1.5% of US patents are ever litigated, and only 0.1% of patents are ever litigated to trial.[42] All litigation could be viewed as a failure to settle. Those that finally reach a court decision are extremely selected: both sides must have a strong belief that the expected value of the

---

the invention's commercial value. In a more complicated model where the application's patentability plays a role in both the USPTO's decision to grant and the applicant's decision to persist, the relationship between the examiner's search effort and patent strength may not be as straightforward. A patent with a higher search effort has two competing implications: (1) the examiner perceives it to be less patentable, but (2) it survived a more rigorous examination with more prior art being reviewed. The results from a more complicated model might be ambiguous.

[41] See Graham et al. (2005) for a detailed discussion and an interesting comparison of USPTO re-examination and EPO opposition. Also, see Merges (1999).

[42] Lemley and Shapiro (2005), p75.

outcome is high; otherwise they would have settled before the trial to save millions of dollars of cost.

In this paper, we instead look at the cases where an inventor files applications, with the same priority date, in multiple patent offices.[43] We study a sample of 22,300 US inventions[44] that filed and *obtained* US patents at the USPTO, and *also* filed applications, through non-PCT filings, at the European Patent Office (EPO) and the Japanese Patent Office (JPO).[45] The US patents in our sample cover 30 distinct technology fields,[46] and their USPTO application years range from 1990 through 1995.

In this paper, we focus on the EPO application outcomes of the US patents in the sample, because the application process at the JPO has its own peculiarities that complicate the issues.[47] Since the EPO applies patentability standards broadly similar to that of the USPTO,[48,49] and the independent EPO examination could be considered as a second trial of the patentability of an invention claimed in a US patent, and EPO application outcomes might betray the strength of the patent.[50] Moreover, since the EPO has been widely viewed as

---

43 Studies that make use of international patenting include Graham et al. (2002) who study US re-examination and EPO opposition process by matching EPO patents to their "equivalent" US patents, Graham and Harhoff (2006) studying a set of litigated US patents and their "equivalent" EPO applications, and Jensen et al. (2006) who document the disharmony in application outcomes by international patenting offices.

44 These inventions are defined as US inventions because they have only US inventors and filed applications at the USPTO first.

45 Our data is a subset of the larger dataset compiled by Jensen et al. (2006). See Jensen et al. (2006) for a detailed description of their dataset and the variables in the data. This study focuses on US inventions only, because they provide a clear timeline of the application process both at the USPTO and the EPO, discussed in subsection 4.2.

46 The 30 technology fields are Office of Science and Technology (OST) technology groups, following Jensen et al. (2006). See Office of Science and Technology, Department of Trade and Industry, United Kingdom classification. See http://www.ipaustralia.gov.au/about/statistics.shtml#patents for details of the classification system. We also use the 6 Technology Categories and 36 Sub-Categories in Hall et al. (2001) to categorize technology fields, and the results remain similar.

47 For instance, an applicant could wait up to seven years to request an examination, until which point the application just sits at the JPO. It is very likely that what happens at the JPO, to a large degree, depends on what happened at the USPTO and the EPO. Studying the interactions among application processes at the Triadic patent office might be an interesting research line in its own right, which we shall pursue in the future.

48 The Agreement on Trade Related Aspects of Intellectual Property (TRIPS) require that all signatories to the agreement apply the criteria of novelty, non-obviousness and utility to determine whether an invention is eligible for a patent (TRIPS Article 27). Specifically, EPO patents are issued for inventions that are novel, mark an inventive step, are commercially applicable, and are not excluded from patentability for other reasons (Article 52 EPC).

49 There remain certain differences between the USPTO and the EPO, mostly in patentable subject matter. For example, software and genes are considered as subject matters that are less patentable at the EPO.

50 An analogy is two independent blood tests at two hospitals for the cholesterol level.

having more rigorous examinations and better performance than the USPTO,[51] an application related to a stronger US patents is more likely to survive at the EPO. In both cases, the outcome of an EPO application might signal the strength (validity) of the corresponding US patent.

Note that the US patents in our sample are also selected in the sense that (1) inventions in our sample sought patent protection in both the USPTO and the EPO through a non-PCT filing. They might differ from those seeking only a US patent, or from those filing a PCT application, in regard to potential commercial value or patentability; and (2) the sample does not include inventions associated with applications that were filed with the USPTO, but rejected.[52, 53] Given these sample selection phenomena, we should use caution in interpreting our findings.

4.2 The EPO application process

For a US inventor who files applications in both the USPTO and the EPO, the time line of the application process is illustrated in Figure 1. The applicant usually files an application at the USPTO first, where an examination automatically ensues, needing no further request from the applicant. Within one year from the US filing date, the applicant needs to file with the EPO in order to claim the US priority date. Upon receiving the application, a centralized EPO search office in The Hague, Netherlands, conducts search for prior art, writes a search report that cites the relevant prior art that the search found, and publishes the search report.[54] Within 6 months after the search report is published, the applicant must decide whether to file a request for examination and pay the examination fee.

---

[51] The grant rate for the USPTO in 1993-1998, corrected for continuing applications, ranges from 80% to 97%. In contrast, the grant rates for the EPO and the JPO from 1995-1999 (averaged) are 67% and 64%, respectively. (Quillen and Webster (2001)).

[52] Before 2000, the USPTO only published issued patents and did not publish applications that were rejected by the USPTO. Thus, we do not have information about rejected applications filed at the USPTO before 2000.

[53] The second selection issue might not be significant, given estimates of the grant rates at the USPTO based on original applications could be as high as 95% (Quillen and Webster, 2001). Even with a lower bound of the USPTO grant rate of 75-80%, the second selection problem might not be serious, as applications that are Non-PCT filings may have a higher grant rate than the general USPTO applications.

[54] The search report is published either together with the publication of the application if the report is available by the due publication date for the application (18 months after the claimed priority date), or alone otherwise.

Upon receiving the request, an EPO examiner (different from the searcher) starts the examination of its patentability. If no request is made, the application is deemed as withdrawn by the EPO. Starting from the third year from the EPO filing date, the applicant must pay annual fees to keep the application alive until the application is disposed of at the EPO. The applicant may decide to withdraw the application at any time during the EPO examination. Meanwhile, the issuance of the US patent can occur at any time during the EPO examination.

There are three EPO application outcomes of interest: withdrawn by applicants, granted or rejected by the EPO.[55] These are sequential events: an applicant first decides whether to withdraw her application and, conditional on a non-withdrawal, the decision of grant or rejection by the EPO is then observed. Figure 2 illustrates the sequential events associated with an application at the EPO.

As shown in Table 1, overall only 60.2% of the US patents in our sample have a corresponding application granted at the EPO; 28.3% withdrawn and 5.8% rejected, respectively. The EPO grant rates vary significantly across different technology fields, ranging from 44% to 75.6%. The technologies with the top three grant rates are handling printing (75.64%), transport (72.7%), and agricultural food (71.6%). The bottom three technologies are semiconductors (44%), pharmaceuticals (45.4%), and information technology (51.4%). The EPO grant rates also vary by years, between 51.5% to 65.5%.

4.3 Information about prior art obtained and missed by US examiners

To construct a measure of the US examiner' effort in search for prior art for a US patent, we need information about the prior art the examiner actually obtains through his search and the prior art he misses.

With regard to the former, on the front page of a US patent document there is a reference section where prior art is cited by the examiner. An applicant has the "duty to disclose" prior art of which he is aware (though no duty to conduct a thorough search for prior art), therefore the cited prior art on the front page of a US patent may contain both prior art disclosed by the applicant, and prior art obtained by the examiner through the search; these

[55] See Jensen et al. (2006) for a more detailed discussion of categorizing these three EPO application outcomes.

cannot be distinguished for US patents in our sample.[56]

Using all prior art cited on the front page of a US patent to proxy for the prior art that the examiner (rather than the applicant) actually identifies through his search may not be a serious problem for our study, for two reasons. First, it is possible that the examiner would have obtained some applicant-disclosed prior art through the search, had the applicant not disclosed it. In other words, some applicant-disclosed prior art might just be a substitute for what the examiner would have obtained. More importantly, using all cited prior art, instead of the prior art obtained by the examiner induces a bias against our results in Section 4 and renders the results even more significant. There is evidence that applicants care more about applications that are more original and more important, for which applicants tend to conduct a search for prior art and disclose more prior art to the USPTO (Sampat 2005)[57, 58]. Therefore, a "better" application might have more applicant-disclosed prior art that, *ceteris paribus*, would render the resulting patent, which is a *stronger* patent, to have an *understated* share of missed prior art and an *overstated* examiner search effort. This would bias against our findings in Section 4 that a stronger patent has a higher share of missed prior art. In other words, if we had information about the prior art obtained by examiners through their search, and used it to construct the variable (the share of missed prior art), our results would have been even stronger and more significant.

Regarding prior art missed by US examiners during their searches, it is very difficult and expensive, if not impossible, to get precise information about missed prior art for a large sample of US patents. There is no way, other than reading patent claims in detail and applying perfect judgment, to identify exactly which prior art is *relevant and missing*. Patent lawyers

---

[56] Starting from 2001, applicant-referenced prior art and examiner-referenced prior art are distinguished in a US patent. But the patents in our sample were filed at the USPTO during the period of 1990-1995.

[57] Sampat (2005) discusses the incentives and disincentives for applicants to search for prior art and disclose to the USPTO. He finds that, after controlling for technology fields, patent applicants devote more effort to identifying prior art for more technologically and commercially valuable inventions, measured by forward citations (citations cited by subsequent patents), 4th year patent renewal and patent family size. He also finds that in so-called "complex product" industries (i.e. electronics, computers and telecommunication technologies), where patenting is mostly for "strategic" purposes of preserving freedom to practice and used as bargaining chips in cross-licensing and where firms care more about the quantity of patents than about obtaining "quality" patents, applicants are less likely to search for prior art.

[58] The findings in Sampat (2005) are consistent with our informal discussions with some ex patent lawyers. Large companies tend to file thousands of patent applications at the USPTO each year, for which they tend to disclose no prior art to the USPTO. Alcacer and Gittleman (2004) study a random sample of 1,500 US patents over the period 2001-2003 and find that 40% of them have no prior art cited by applicants.

try to do this in patent litigation, for which they charge high fees. As an alternative, our study relies on a sophisticated Latent Semantic Analysis (LSA), also known as Natural Language Processing (NLP) analysis, to determine how closely related other patents are to a given patent (hereafter, *root patent*) and thus identify the missed prior patents of the root patent.[59] When a prior patent satisfies a *particular threshold of linguistic similarity* to the root patent and is not cited by the root patent, it will be flagged as a missed prior patent of the root patent.[60, 61] The LSA analysis used in the paper was kindly provided by M-CAM, a Charlottesville, VA-based patent analysis firm, which employed its own NLP algorithm to identify missed prior patents for each US patent in our sample.[62, 63]

For each patent *i* in the sample, in addition to "M-CAM missed prior patents" (MPP), the M-CAM analysis also provides two more pieces of information that we use in the study. The first piece is "M-CAM linguistically linked prior patents" (LLPP), all the prior patents that the M-CAM Latent Semantic Analysis identifies as linguistically linked to, *but not cited by,* the patent *i*. LLPPs are located by the M-CAM algorithm using a much lower threshold of linguistic similarity than MPPs and used as a proxy for difficulty of finding prior patents.[64]

---

[59] For an explanation and application of the NLP analysis, visit http://www.cognition.com/info/how.html. In addition to M-CAM (www.m-cam.com) that does a NLP analysis for the US patents in our sample, there are other companies whose business is based on NLP. Among them is Cognition Technologies, Inc. (www.cognition.com).

[60] Note that the M-CAM analysis only searches patent documents to retrieve what it deems as "missed prior patents", and we don't have data on other types of missed prior art, such as missed journal articles. In this paper, we use only information about cited prior patents and M-CAM missed prior patents to measure US examiners' search effort. It is widely viewed and empirically confirmed that US examiners' searches are primarily focused on patent documents, particularly on US patents (Sampat 2005). Thus, measuring an examiner' search effort using the number of cited prior patents and the number of missed prior patents appears reasonable, if not ideal.

[61] The paper reports the results with examiners' search effort measured by the number of cited prior US *patents* and the number of missed prior US *patents*. Using all the cited and missed prior patents (US, EP, Japanese and other national patents) might have a potential double counting problem, because one prior invention might be protected in multiple jurisdictions. The results with the latter measure of examiners' search effort, not reported here, are similar.

[62] The analysis algorithm of M-CAM is a trade secret, and detailed information about the algorithm is therefore unavailable. However, the firm indicates that the algorithm incorporates elements of latent semantic analysis as well as employing patent-specific bibliographical information to determine relationships between patents.

[63] Of course, the M-CAM algorithm is not comparable to a legal and technical assessment of what are missed prior patents. However, on average, and over a large data set, the algorithm could provide useful information about technically related but uncited prior patents.

[64] Analogous to those hundreds of links that appear if we do a Google search using some keywords, these linguistically linked prior patents (*LLPP*) can be considered as the pool of potential prior patents that the US examiner gets with a first round of search using some keywords, from which the examiner then has to figure out the "true" prior patents.

The second piece of information relates to the subsequent patents that might have built on root patent *i* and should have cited patent *i* but did not. We call them "M-CAM unciting subsequent patents," as opposed to those subsequent patents that cite the root patent (hereafter citing subsequent patents). We combine a patent's "citing subsequent patents" and "M-CAM unciting subsequent patents" together and use these "total subsequent patents" as a possibly more accurate measure for the importance and value of the patent. [65] Appendix 1 gives a more comprehensive description of the information obtained from the M-CAM analysis.

We conduct a case study to check how good the M-CAM LSA analysis is, i.e., how likely the M-CAM missed prior patents are the "true" missed prior patents. Listed in Table 2 are six high-profile patents that have been revoked by the USPTO after re-examinations, thanks to the validity challenges brought by PubPat, a non-profit organization representing the public interest and specializing in challenging undeserved patents that are both economically and socially significant.[66] In three out of six cases, the true missed prior patents that were used by PubPat to invalidate those undeserved patents are included in the "M-CAM missed prior patents". In five out of six cases, the true invalidating prior patents are included in the broader class, "M-CAM linguistically linked prior patents." Thus, we have some confidence that M-CAM algorithm does a reasonably good job in locating missed prior patents.

Figure 3 shows the histograms of the number of cited prior patents and the number of M-CAM missed prior patents for the US patents in the sample. Both distributions are concentrated in the range of 1-50 and show significant levels of dispersion, suggesting that the M-CAM algorithm achieves a degree of discrimination' it does not mechanically retrieve a roughly similar number of prior patents for each US patent in the sample.

4.4 Measurement for US examiner's search effort: share of missed prior art

For each US patent *i* in our sample, based on the number of the cited prior patents *(CPPi)* on the front page and the number of M-CAM missed prior patents (MPP_i), we construct a variable, *share of missed prior patents* ($SMPP_i$), where $SMPP_i$=

[65] It has been suggested that if a patent has more citing subsequent patents and/or the citing subsequent patents spread over more diverse technology fields, the patent is more important and more valuable. See Hall, Jaffe and Trajtenberg (2001) and cited references there.

[66] For more information about PubPat and those revoked patents, visit http://www.pubpat.org/.

$MPP_i/(MPP_i+CPP_i)$ and (MPP_i+CPP_i) is the number of total prior patents. We are interested in whether a higher $SMPP_i$ suggests a stronger or weaker patent.

A key advantage of using "M-CAM missed prior patents" is that, for the US patents in our sample, this information is observed only by us and was not known by either applicants or examiners in any patent offices. Thus, it has no *direct* impact on decisions by applicants and patent examiners.[67] However, note that the set of M-CAM missed prior patents is not the "true" set of missed prior patents in a definitive sense. The constructed variable, $SMPP_i$ is a noisy measure of US examiner's search effort, which could lead to a potential attenuation problem. This potential attenuation problem, caused by a noisy SMPP, might lead to understatement of the significance of the results, reported later.

Figure 4 provides graphic evidence that a higher SMPP suggests a stronger patent. It compares mean differences in SMPPs between US patents with different EPO application outcomes. We consider three comparisons: (1) the US patents that were not withdrawn at the EPO versus those withdrawn, (2) those granted by the EPO versus those rejected, conditional on non-withdrawal, and (3) those successful (granted) at the EPO versus those failed (either rejected or withdrawn).

Panel A of Figure 4 shows the mean differences in SMPPs for the three comparisons for US patents in each of the 30 technology fields. A solid circle indicates a significant difference from zero at 5% level, while a hollow circle indicates an insignificant difference. In 7 out of 30 technology fields,[68] the US patents that were not withdrawn at the EPO have a significantly *higher* mean value in SMPP than those withdrawn; and in no fields, did those not withdrawn have a significantly lower mean SMPP relative to those withdrawn. The mean differences in SMPP between those granted and those rejected (conditional on a non-withdrawal) are significantly positive for 6 technology fields[69] and insignificant for other technologies. In 10 out of the 30 technology fields,[70] the differences in the average SMPP

---

[67] M-CAM started its patent analysis business in 1999.

[68] Those technology fields include telecommunications, optics, organic fine chemicals, macromolecular polymer, surfaces coatings, transport and nuclear engineering.

[69] The fields include electrical devices, medical engineering, biotechnology, materials metallurgy, space technology weapons, and civil engineering building mining.

[70] The fields include electrical devices, telecommunications, medical engineering, organic fine chemicals, macromolecular polymer, materials metallurgy, surfaces coatings, transport, nuclear engineering, and civil engineering building mining.

between the success (granted) group and the failure (rejected or withdrawn) group are significantly positive, and in no field is a significantly negative difference observed.

Panel B of Figure 4 shows the mean differences in SMPPs for the three comparisons for US patents with each of the six USPTO application years (1990-1995). In 5 out of the 6 years, US patents that are not withdrawn at the EPO have a significantly higher mean SMPP than those withdrawn; and only for year 1992, the difference is insignificantly positive. The differences in average SMPP for those granted and those rejected by the EPO (conditional on non-withdrawal) are significantly positive in two years (1990 and 1994), insignificantly positive in three years (1991, 1993 and 1995), and insignificantly negative only in 1992. In five out of the six years there are significantly positive mean differences in SMPP between the success (granted by the EPO) and failure (withdrawn or rejected) groups; the mean difference is insignificantly positive only in 1992.

4.5 Control variables

With regard to control variables, we use the number of "M-CAM linguistically linked prior patents" (LLPP) as a proxy for difficulty of finding relevant prior inventions. We also control for other characteristics of an invention that may influence applicants' behavior or have impacts on EPO decisions. For instance, the higher commercial value an applicant believes his invention has, the more likely he is to be persistent and spend more money to get a European patent. As a result, the probabilities of not being withdrawn and being granted by the EPO (conditional on a non-withdrawal) might be higher if commercial value is higher. Specifically, for each US patent in the sample, we include the following control variables:[71]

(1) The number of claims in the US patent,

(2) The number of US classifications the US patent belongs to,

(3) The number of inventors listed in the US patent,

(4) The number of assignees listed in the US patent,

(5) The number of total prior patents (both cited and missed prior patents),

[71] We use "total prior patents" and "total subsequent patents" to construct control variables in the list because they, unlike "cited prior patents" and "citing subsequent patents" that are decided by examiners, are exogenous. As checks on the robustness of our results, we also use "cited prior patents" and "citing subsequent patents" to construct control variables and the results are similar.

(6) The number of US classifications that these prior patents belong to,[72]

(7) The number of total subsequent patents (both citing and unciting subsequent patents).

(8) The number of US classifications these subsequent patents belong to,[73]

(9) Three potential indicators of the *technology stage* of the US patent, i.e., whether it is at an early or late stage in its own technological trajectory: *Innovation stage* (the number of total prior patents over the number of total subsequent patents), *Lag to total prior patents* (the length of the period between the average issue date of the total prior patents and the issue date of the US patent), and *Lag of total subsequent patents* (the lag between the issue date of the US patent and the average issue date of the total subsequent patents). If the first two indicators are larger, the patent might be in a later stage, while a larger third indicator might suggest an earlier invention.

5. Empirical Strategies and Results

In this section, we first describe our empirical strategies used in testing the relationship between SMPP and patent strength. We then present the empirical results that, for US patents with a higher SMPP, their related EPO applications are more likely not to be withdrawn by applicants and, conditional on non-withdrawal, more likely to receive patents at the EPO. Our interpretation of these results is that a higher SMPP indicates a stronger patent. We then test three alternative explanations for the finding that a US patent with a higher SMPP is more likely to succeed at the EPO and find them implausible. Finally, we have a few robustness checks and further verify our results.

5.1 Empirical strategies

A patent *i*'s share of missed prior patents, $SMPP_i$, is not only a function of the

---

[72] Also, following Hall et al. (2001), we construct an originality variable based on the US classifications that the prior patents belong to. We also use the originality variable that is constructed by Hall et al. (2001). In both cases, the results still hold.

[73] We also, following Hall et al. (2001), construct a generality variable based on the US classifications that the subsequent patents belong to, and the results still hold. We did not use the generality variable in Hall et al. (2001), because the US patents in our sample were issued from mid 1990s onward and would have many subsequent patents with an issue year beyond 1999, which are not included in Hall et al. (2001).

patentability of an application, it also depends on the characteristics of the US examiner who did the examination. It is well known that there exists significant heterogeneity in many aspects across US examiners. Indeed Cockburn et al. (2004) state that "there may be as many patent offices as there are patent examiners." To control for the potential heterogeneity in search effort across US examiners, we collect the information about the names of primary examiners and assistant examiners (if any) for each of the US patents in the sample. An ANOVA test confirms that the variable SMPP differs systematically for patents examined by different primary examiners. To control for the heterogeneous search effort across primary examiners, we use a *US examiner* fixed effect for each primary examiner, which allows us to use variations in SMPPs within patents examined by the same US examiner to identify the effect of SMPP.[74] The intuition is that if a patent's SMPP is higher relative to the other patents examined by the same examiner, then the examiner might have thought this patent is more patentable and made less of a search effort. We also use a dummy variable for whether a US patent has an assistant examiner to control for the potential difference between US patents that are examined by both a primary examiner and an assistant examiner, and those examined by a primary examiner alone

A patent's SMPP also depends on its technology field. Figure 4 plots the average SMPP and EPO grant rate for each of the 30 technologies, with the average SMPPs in an ascending order. We can see that the average SMPP differs from one technology to another. But there is no clear correlation between the average SMPPs and the EPO grant rates. US patents in medical engineering have the highest average SMPP and patents in biotechnology the lowest, but the EPO grant rates for patents in medical engineering and biotechnology are about the same and rank in the middle among these 30 technologies. To control for the heterogeneity in SMPP across technology fields, dummy variables for each of the 30 technologies are used. We also use dummy variables for different USPTO application years to control for unobserved time-varying factors at the USPTO that may impact SMPP.

[74] We focus on primary examiners in our empirical specifications, as an assistant examiner does not have signatory authority in issuing patents. 50.1% of the patents in the sample does not have an assistant examiner.

We use two empirical strategies in the paper. In the first strategy, we include US patents whose examiners have at least *10* patents in our sample[75] and run the following unbalanced panel data model with *examiner* fixed effects:

(A) $$Y_{ieft} = \alpha_1 * SMPP_i + \beta_1 * X_i + \gamma_e + \delta_f + \eta_t + \mu_{ieft}$$

Where Y_{ieft} is the EPO application outcome for a US patent *i* that is examined by a US examiner *e*, in technology field *f* and with a USPTO application year *t; $SMPP_i$* is the share of missed prior patents for the patent *i; X_i*'s are control variables for the patent *i* that influence its EPO application outcome, described in Section 4.6; γ_e is the examiner fixed effect; and δ_f and η_t are dummies for the 30 technology fields and six USPTO application years, respectively.

In the second strategy, we use *examiner by technology by year* fixed effects to control for unobserved time-varying factors whose impacts may differ across different examiners and/or different technology fields. Over the course of six years (1990-1995) some US examiners might gain more experience than other examiners, and some technology fields might evolve more dramatically than other field.[76] To control for these unobserved examiner-specific and technology-specific factors that vary over time, we include US patents whose examiners have at least *4* patents in the sample that are in the same technology field and with a same USPTO application year, and run an unbalanced panel data model with *examiner by technology by year* fixed effects, θ_{eft}:

(B) $$Y_{ieft} = \alpha_2 * SMPP_i + \beta_2 * X_i + \theta_{eft} + \nu_{ieft}$$

Essentially, we identify the effect of SMPP on EPO application outcomes using variations in SMPPs within US patents that are examined by the same examiner *e*, in the same technology field *f* and filed at the USPTO in the same year *t*. Since in each year a US

[75] We also try different cut-off points and use US patents whose examiners have at least 5, 15 and 20 patents in the sample, respectively. The results remain similar.

[76] For instance, in early 1990s nanotechnology was a very new field and US examiners did not know much about it and thus cited little prior art. In time, examiners gained more knowledge about this technology and the amount of cited prior art increased gradually.

examiner might not have multiple patents in our sample that are in the same technology field, the second strategy significantly reduces the size of the analytic sample, leading to larger standard errors for the estimated coefficients. It is a more robust and convincing strategy and a nice complement to the first empirical strategy.

For both strategies, a linear probability model is used when the dependent variable is binary, which is the case for most of the regressions in the study (e.g., applicants' decisions not to withdraw at the EPO, the EPO's decisions to grant or reject, and so on).

Panel A of Table 3 shows the summary statistics for the whole sample. Panel B shows the summary for the analytic sample used in the first empirical strategy (A) that involves *examiner* fixed effects. Panel C reveals the summary statistics for the analytic sample in the second empirical strategy (B) using *examiner by technology by year* fixed effects. There is no systematic difference across the three panels.

5.2 EPO application outcomes

The regression results confirm the message revealed graphically in Figure 5. Table 4 and 5 show the regression results regarding the EPO application outcomes. The empirical strategy in Table 4 involves an unbalanced panel data model with *examiner* fixed effects, and the strategy in Table 5 implements an unbalanced panel data with *examiner by technology by year* fixed effects. Various specifications are tested in both strategies.

Panel A of Table 4 studies the probability of non-withdrawal at the EPO, with the dependent variable being 1 if not being withdrawn and 0 otherwise. The estimated coefficients for the share of missed prior patents (SMPP) are all significantly positive at the 1% level. An application related to a US patent with a higher SMPP is more likely not to be withdrawn at the EPO.

The result is consistent with the hypothesis that US examiners distinguish good applications from bad ones and put more effort in search for prior art for those deemed as less patentable. Examiners are not "rationally ignorant" of applications' patentability.

The coefficients for the variable *LLPP* (logarithm of the number of linguistically linked prior patents), a possible indicator for difficulty of finding relevant priors, are not significant. Neither are the coefficients for whether a US patent had an assistant examiner. A

US patent in a later technology stage, measured by the variable *Innovation stage*, is more likely to be withdrawn at the EPO. If a US patent has more assignees, it is less likely to be withdrawn at the EPO, possibly because having more assignees implies a higher commercial value, or it is harder for multiple assignees to agree on withdrawal.

Panel B of Table 4 shows the results for the EPO's decision to grant or reject an application, conditional on that the applicant persists at the EPO and does not withdraw it. The dependent variable is 1 if the EPO grants and 0 if rejects. Across various specifications the estimated coefficients for SMPP are all significantly positive, suggesting that a US patent with a higher SMPP has more probability of being granted by the EPO. Note that this subset of US patents has already survived a first round of filtering by applicants' decisions of whether to withdraw and thus, on average, have higher SMPPs than those withdrawn. Still, the variable SMPP has a significantly positive correlation with the EPO's decision to grant. For control variables, if a US patent has a higher number of total subsequent patents, an indicator of the patent's importance or value, it is more likely to be granted by the EPO, conditional on a non-withdrawal. This result further confirms our main hypothesis that a US examiner's search effort is dependent on his perception of the patentability of an application and his perception is quite consistent with that of his EPO counterparts.

In Panel C of Table 4, we group together those US patents whose corresponding EPO applications were either withdrawn by applicants or rejected by the EPO as the failure group, as opposed to the success group composed of the US patents that were granted. The variable, SMPP, exhibits a significantly positive effect on the probability of being successful at the EPO. An increase in SMPP by 0.5 (roughly from the 25th percentile to the 75^{th} percentile in SMPPs for the US patents in the sample) will increase the EPO success rate by about 5 percentage points. Given the EPO success rate for the analytical sample is around 65% and the variable SMPP is constructed using M-CAM uncited prior patents, probably a very *noisy* proxy for US examiner's search effort, this is a strong result.

The number of assignees is positively correlated with the likelihood of success at the EPO. A US patent that is in a later stage of its technology trajectory, indicated by a larger *Innovation stage*, is more likely to fail at the EPO. If the total prior patents of a US patent are more dispersed in different technologies (covering more US classifications), it is less likely for the patent to be successful at the EPO.

The regressions in Table 5 involve *examiner by technology by year* fixed. The results in Table 5 are similar to those in Table 4, but the standard errors of the estimates are larger, as the size of the analytic sample is smaller. These results further confirm that US examiners can, by and large, make a good assessment of the patentability of an application, and conduct their searches for prior art accordingly.

5.2 Applicants' request for examination after EPO search report

To confirm our findings, we investigate an applicant's decision on whether to request an examination after she receives an EPO search report. As described in Section 4.2, an EPO search report cites relevant prior art that an EPO searcher identifies for an EPO application, and it is the first important information about the patentability that the applicant learns about from the EPO. Within 6 months after receiving the EPO search report, the applicant must decide whether to request an examination. In our sample, 94% did so with an average lag between requesting and receiving search reports being 4.3 months. For the remaining 6%, their applications were deemed as withdrawn.

Table 6 shows the results. Panel A in Table 6 implements an *examiner* fixed effect model and Panel B an *examiner by technology by year* fixed effect model. The results from these two strategies are similar. Perhaps due to a smaller sample size, some coefficients in Panel B are less significant than their counterparts in Panel A, though the magnitudes are close. Columns 1 and 3 show that the variable, SMPP, has significant explanatory power in the likelihood of requesting an examination at the EPO after the applicant receives the EPO search report. Furthermore, if an applicant is to request an examination, she does so *sooner* if the US patent has a higher SMPP, as shown in Columns 2 and 4 in Table 6. If a US patent has a higher number of US classifications, a possible indicator of the complexity of the invention, the applicant takes a longer time to request an examination in the EPO. A higher number of assignees leads to a faster request, suggesting that the variable is more of an indicator of commercial value than of the difficulty of reaching consensus.

5.3 Testing alternative rationales

Thus far, the results have shown that for a US patent with a higher share of missed prior patents (SMPP), its corresponding EPO application is more likely not to be withdrawn and, conditional on non-withdrawal, more likely to be granted by the EPO. Our interpretation is that when an invention is more patentable (more novel and non-obvious), the US examiner would make less of an effort in searching for prior art and the resulting US patent would have a higher SMPP. Since the invention is more patentable, it will more likely succeed at the EPO. Therefore, the variable SMPP, reflecting the patentability (strength) of an invention, has significant explanatory power regarding the EPO application outcome.

In this subsection, we also address several alternative explanations for the finding that a US patent with a higher SMPP is more likely to succeed at the EPO. We argue that these alternative stories are implausible.

Alternative story 1: information flow

The first alternative explanation tells a story of information flow from the USPTO to the EPO. Suppose that during their search for prior art, EPO searchers rely heavily on what US examiners have found. For instance, in a extremely hypothetical case then EPO searcher just copies the prior art that are *cited* by US examiners in issued US patents. In this case, a higher SMPP, caused by whatever reasons at the USPTO, may be transferred to the EPO. As a result, the EPO examiner would miss a higher portion of prior art when doing their examination, and therefore, be more likely to grant a patent. In this story, the variable SMPP has nothing to do with US examiner's search effort.

To test this explanation, we look at a subset of US patents, for which EPO search reports had been published ***before*** the US patents were issued. This subsample covers for 41.5% of the US patents in our sample. For them, it was impossible for EPO searchers to know what prior art US examiners would cite. Table 7 shows the results, for both an *examiner* fixed effect model and an *examiner by technology by year* fixed effect model. The findings still hold that a US patent with a higher SMPP is more likely not to be withdrawn and more likely to succeed at the EPO. Therefore, the story of information flow from the USPTO to the EPO is unpersuasive.

Alternative story 2: difficulty of finding prior art

The second alternative explanation is that it might be inherently more difficult for some inventions to find prior art, even after we control for technology fields, the number of total prior art, and the number of linguistically linked prior art (LLPP). If this is the case, the variable SMPP might just pick up the effect of such difficulty of finding prior art. An invention for which it is more difficult to find its prior art might have a higher SMPP at the USPTO and, when it goes to the EPO, the EPO searcher might miss a larger portion of its prior art, causing the EPO examiner to be more likely to grant.

In Table 8, we show that a US patent with a *highe*r SMPP actually had a *shorter* search interval in the EPO search office. This contradicts the story that it is the difficulty of finding prior art that causes a higher SMPP. The result seems to suggest that EPO searchers have the same search strategy as US examiners: for a more novel/non-obvious invention, they tend to search less, resulting in a shorter search interval.

Columns 1 and 3 of Table 8 investigate the relationship between SMPP and the EPO search interval, measured by the time between the EPO filing date and the publication date of the search report.[77] The coefficients for SMPP are significantly negative. The number of linguistically linked prior patents (LLPP) is positively correlated with the EPO search length, suggesting that, all things equal, it takes EPO searchers more time to go through a larger pool of potential prior patents to figure out which are the relevant prior art. If a US patent has more claims, its EPO application may have more claims as well; and that leads to a longer EPO search time. A US patent in a later technology stage, indicated by a longer *Lag to total prior patents*, has a shorter EPO search process, possibly because EPO searchers are more familiar with the technology field and thus the search process is easier and shorter.

Columns 2 and 4 of Table 8 investigate the average search time spent on each piece of prior art cited in the search report.[78] The variable SMPP is also negatively correlated with the search interval per citation, further suggesting that a higher SMPP is not caused by the inherent difficulty of finding prior art.

---

[77] The average EPO search interval is 1.05 years.

[78] The Jensen et al. (2006) dataset contains a field about the amount of cited prior art in EPO search reports. We construct a variable, search interval per cited prior art, by dividing the total search time by the number of references cited in the search report. The average search interval per cited prior art is 0.36 years.

Alternative story 3: a novel invention has fewer prior art

The third alternative explanation is based on a widely accepted notion that a more original (novel and non-obvious) invention tends to have less prior art. Suppose that the US examiner actually obtained *all* relevant prior art and there is no so-called “missed prior art” at all and that, for each US patent in the sample, what the M-CAM linguistic/semantic matching algorithm gives us, is roughly a similar number of prior patents that only bear some linguistic similarity to the root patent. In this case, we would have artificially constructed a higher SMPP for a more novel/non-obvious US patent, even though there are no “missed prior patents”.

We don’t think this story as plausible, on two grounds. First, the histogram in Figure 3 shows that the number of *missed* prior patents has a significant degree of dispersion and is not concentrated in a narrow range. Secondly, we test whether the number of *missed* prior patents is correlated with EPO application outcomes, particularly when the number of *cited* prior patents is controlled for. According to this third alternative explanation, missed prior patents should bear no information about the invention’s patentability. Part (a) of Table 9 shows the effects of the number of *cited* prior patents and the number of *missed* prior patents on EPO application outcomes, with an *examiner* fixed effect model. The number of *missed* prior patents is significantly and positively correlated with the probability of an EPO success, both when excluding and including the number of *cited* prior patents in the regressions. Thus, the number of missed prior patents does contain significant information about the patentability, rendering the alternative story implausible.

The number of *cited* prior patents is negatively correlated with the likelihood of an EPO success. When a US examiner puts more effort in search for prior art for an application deemed as less patentable, he obtains and cites more prior art. Interestingly, the effects for the number of missed prior patents and the number of cited prior patents are both greater when the two variables are both included in the regressions, suggesting that the ratio, the share of missed prior patents (SMPP), is a better measure of the US examiner’s search effort than either the number of missed prior patents or the number of missed prior patents alone.

Part (b) of Table 9 involves a *examiner by technology by year* fixed effect model. The results are similar in magnitude to those in Part (a). Some coefficients become less significant because of larger standard errors, possibly due to a smaller sample size.

After testing these alternative explanations, we further verify the hypothesis that a higher SMPP is caused by less effort in search for prior art by the US examiner, who makes such a decision according to the patentability of the application.

5.3 Robustness checks

We also did three robustness checks, two involving specifications that employ different fixed effects in the panel data model and another testing the outcomes of the related JPO applications. All these robustness checks further verify the result that a US patent with a higher SMPP is a stronger patent.

US examiner by patent assignee fixed effect model

We also implement a *US examiner by patent assignee* fixed effect model. A *US examiner by patent assignee* fixed effect allows us to look at US patents that are filed by the same US firm and examined by the same US examiner and control for unobserved heterogeneities in application strategies by different applicants (large firms versus small firms) and/or heterogeneity in examiners' responses to different applicants:

(C) $$Y_{ieaft} = \alpha_3 * SMPP_i + \beta_3 * X_i + \gamma_{ea} + \delta_f + \eta_t + \nu_{ieaft}$$

Y_{ieaft} is the EPO application outcome for a US patent *i* that is assigned to an assignee *a*, examined by a US examiner *e*, in technology field *f* and with a USPTO application year *t;* γ_{ea} is the *US examiner by patent assignee* fixed effect; δ_f and η_t are dummies controlling for technology fields and USPTO application years, respectively.

Table 10 shows the results of various specifications, all using a *US examiner by US assignee* fixed effect model. The sample size is much smaller because many pairs of US examiner-assignee do not have multiple patents in our sample. In all the specifications, the coefficients for the share of missed prior patent, SMPP, are all significantly positive with regard to the probabilities of not being withdrawn, of being granted conditional on non-withdrawal, and of being successful at the EPO.

Primary examiner by secondary examiner fixed effects

So far, we have treated patents with the same primary examiner to have the same examiner fixed effect and only used a dummy variable to control for whether there is a secondary examiner. Our reasoning is that it is primary examiners who supervise secondary examiners, make decisions and sign off grant or rejection decisions. But one might argue that since it is secondary examiners who do actual work, there might be fundamental difference among, for instance, a patent with a primary examiner only, a patent with the same primary examiner and a secondary examiner, and a patent with the same primary examiner and another secondary examiner.

Table 11 investigates whether using *primary by secondary examiners* fixed effects change the results. In Part (a) of Table 11, we assign a different examiner ID to each of the standing-alone primary examiners and primary-secondary examiner pairs, and then apply either *examiner* fixed effects (Panel A of Table 11) or *examiner by technology by year* fixed effects (Panel B of Table 11) in the analysis. In Part (b) we look at the subset of patents that have both a primary and secondary examiners, assign a different examiner ID to each pair of primary-secondary examiners, and then apply either *examiner* fixed effects (Panel A) or *examiner by technology by year* fixed effects (Panel B) in the analysis. Table 11 shows that in all these specifications, the results still holds that SMPP significantly predict the probability of non-withdrawal and the probability of being successful at the EPO.

Outcomes of related JPO applications

The applicants in our sample also filed related applications at the JPO, within one year after they filed with the USPTO. Unlike filings at the USPTO or at the EPO, where an examination (at the USPTO) or a search for prior art (at the EPO) automatically ensue, filing at the JPO does not trigger any action by the JPO. An applicant can wait up to seven years to request an examination, until which point the JPO takes no action at all. Therefore, an applicant's decision on whether to submit a request with the JPO, and subsequent decisions on whether to withdraw might be, to a large extent, related to what happened in the USPTO and in the EPO.[79]

In this paper, we run a preliminary reduced form regression to test whether the share of missed prior patents, SMPP, predicts JPO application outcomes.[80] Table 12 shows the results, both for an *examiner* fixed effect model and an *examiner by technology by year* fixed effect model. The coefficients for SMPP are significantly positive in an applicant's decision whether to withdraw at the JPO, though smaller than the coefficients of SMPP in his decision on whether to withdraw at the EPO. This result has two implications. First, it further confirms our main hypothesis that a patent with a higher SMPP is indeed a stronger patent. Second, it seems that, due to various institutional and cultural reasons,[81] the JPO application outcome is a noisier indicator of an invention's patentability than the EPO application outcome.[82] The coefficients for SMPP are not significant in the JPO's decision to grant, conditional on the application not being withdrawn, though still positive. Not surprisingly, the variable SMPP is significant in predicting whether a US patent is successful (granted) or failed (rejected or withdrawn) at the JPO.

6. Policy Implications and Conclusion

In the vast economic literature regarding patents, little has been written about the implication of the behavior of examiners and applicants during patent examination. The process of patent examination has been treated as a "black box," out of which patent applications are either granted or rejected. This paper shows that the patent examination process reveals much richer information than a decision of approval or rejection. We study the

---

[79] The interaction of patent examinations at the triadic patent offices, and its implications for international patenting, are separate issues which we shall study in a separate but related paper.

[80] For the US patens in the sample, 31.1% were withdrawn, 38.9% granted, 16.7% rejected, and 13.2% pending at the JPO.

[81] For example, a US patent application has to be translated to Japanese in order to file at the JPO. Also, prior to 1988, only one independent claim was allowed in a Japanese patent. That rule was changed in 1988 （Sakakikibara and Branstetter (2001). However, Japanese patents, on average, still have fewer independent claims than US and European patents. In 2003, the average number of claims in a Japanese patent was 7, whereas it was 23 for the USPTO and 18 for the EPO.

[82] A simulation shows that in the case of a limited dependent variable, random measurement error in the dependent variable will cause estimated coefficients to be attenuated. To see this, suppose JPO application outcomes are totally random, the coefficient of SMPP would be zero.

effort by US examiners in their searches for prior art. We find that examiners are not "rationally ignorant" of the patentability of patent applications. Examiners can, on average, identify which applications are stronger in two senses: they are more likely to receive a related patent in the Europe; and their US patents are more likely to be renewed by patentees in the USPTO eight year after issuance. For applications that seem weaker, examiners put more effort in search for prior art because they bear the burden of proof of non-patentability.

Our study provides a piece of empirical evidence that the problem of weak patents in the US might be broad and systematic, which, in spite of many anecdotes, has not been empirically tested.[83] Due to lack of compelling quantitative evidence, defenders of the current patent system assert that whatever transition problems the USPTO may have had in such new areas as software and business methods, weak patents will be adequately addressed as the USPTO gains experience and skill in these areas.[84] This study challenges this assertion. Our sample contains all the US patents that were originated from the US (by US inventors) with a USPTO application year between 1990 and 1995 and also applied with the EPO through Non-PCT filings. Among them, almost 35% of their corresponding EPO applications are either withdrawn or rejected. The result that US examiners' search effort has significant effects on EPO application outcomes suggests that if these US examiners worked at the EPO they would also likely reject these applications. They were granted at the USPTO, possibly because US examiners' hands were "tied up" by pro-"customer" rules and procedures. The fact that these applications went through the USPTO but failed at the EPO seems to suggest that the USPTO issues weaker patents.

Furthermore, our study suggests that, in regard to the causes of weak patents in the US, pro-"customer" rules and policies, such as the institutional, procedural and cultural incentives at the USPTO that favor issuing patents, might be more salient than the alleged "rational ignorance" of US examiners. Our results show that US examiners can, in general,

---

[83] The NAS Study (2004) studies the trend in patent strength using three measures: (1) the ratio of invalid to valid patent determinations in infringement law suits; (2) the error rate in USPTO quality assessment review of allowed patent applications; and (3) the rate of claim cancellation or outright patent revocation in reexamination proceedings at the USPTO. These indicators show mixed results. Also, note that measures (1) and (3) are plagued by selection bias.

[84] See, e.g. Edward G. Fiorito, Chair's Bulletin, 2001 A.B.A. Sec. Intellectual Prop. L. Rep. 5, http://www.abanet.org/intelprop/chair/apr01chair.html.

distinguish "good" applications from "bad" ones and search harder for those deemed as less patentable. But there might be too many barriers and disincentives for them to make full use of their knowledge of applications' patentability.

What could be done to address the problem of weak patents? Among proposed measures to improve patent quality, one of the most widely proposed remedies, and also the most intuitively obvious one, is to devote more resources to the USPTO to hire more examiners, provide more technical training for examiners, offer higher salaries to retain senior and experienced examiners, and allocate more time for each application.[85] On the other hand, Lemley (2001) suggests that "spending more time and money weeding out bad patents and strengthening the examination process is not cost effective." Our study suggests that a tradeoff between weeding out weak patents and increasing USPTO expenditure does not necessarily exist at the margin. US examiners, with resource currently available to them, are not "rationally ignorant" of applications' patentability. It seems that, even keeping unchanged the number of examiners, the workload, and the time allocated for each application, the strength of issued patents might be significantly improved by empowering examiners to be able to reject applications that they consider as unpatentable. Measures such as limiting the number of continuations an applicant can file and raising the patentability bar, particularly for non-obviousness, might be effective without incurring much further expenditure at the USPTO.

Finally, our study suggests an alternative view about cited prior art, particularly its implication for patent validity (strength). Scholars have suggested that citing more prior art makes a patent stronger and more likely to stand a validity challenge, because the examination seems to be more thorough and thus it is less likely that prior art exists that will be discovered later and used to invalidate the patent. This argument assumes a non-informative prior art search. Our study, however, indicates examiners, bearing the burden of proof, put more search effort to obtain more and better evidence and to deter applicants from persisting. A higher amount of cited prior art might reflect examiners' assessment of lack of patentability of an application.

---

[85] See NAS Study (2004), FTC Report (2003), and USPTO "The 21st Century Strategic Plan" (Feb, 2003).

Our results in Table 8, consistent with some previous empirical studies,[86] show that for the US patents in our sample, a higher number of *cited* prior patents are positively correlated with the failure at the EPO. A higher amount of cited prior art indicates the weakness of a patent, rather than the survival of a more rigorous examination, partly because issuing a US patent itself does not tell us much about its strength, as the applicant can always persist until the US examiner concedes.

This paper is just our first attempt to look inside the "black box" of patent examination. Future research will use information about office actions by US examiners during examinations, from the USPTO's Patent Application Information Retrieval (PAIR) system, to investigate whether those weak applications were indeed initially rejected but eventually granted by US examiners. Our study opens an interesting research line that studies the behavior of both applicants and examiners, using detailed information about the dynamic progress of the examination processes at multiple patent offices.

[86] See Allison et al. (2004) on the positive relationship between cited prior art and litigation; Allison and Tiller (2003) on the difference in cited prior art between Internet business method patents, for which the weak patent problem is believed to be much more widespread, and the general patents; and Harhoff and Reitzig (2004) on the positive correlation between cited prior art and opposition at the EPO.

References

Alcacer, Juan and Michelle Gittelman, "How Do I Know What You Know? Patent Examiners and the Generation of Patent Citations," New York University, Working Paper, 2004.

Allison, John R. and Mark A. Lemley, "Empirical Evidence on the Validity of Litigated Patents," American Intellectual Property Law Association Quarterly Journal, 26, 185- 228, 2000.

Allison, John R., Mark A. Lemley, Kimberly A. Moore and R. Derek Trunkey, "Valuable Patents," Georgetown Law Journal, 92 (3), 435, 2004.

Allison, John R. and Emerson H. Tiller, "The Business Method Patent Myth," Berkeley Technology Law Journal, 18, 987, 2003.

Choi, Jay Pil, "Live and Let Live: a Tale of Weak Patents," Journal of the European Economic Association, 3, 724-733, 2005.

Cockburn, Iain M., Samuel Kortum and Scott Stern, "Are All Patent Examiners Equal? Examiners, Patent Characteristics, and Litigation Outcomes," in Wesley M. Cohen and Stephen A. Merrill eds, *Patents in the Knowledge-Based Economy, National Academy Press*, Washington, D.C., 19-53, 2003.

Farrell, Joseph and Robert Merges, "Incentives to Challenge and Litigate Patents," Berkeley Technology Law Journal, 19, 943, 2004.

Farrell, Joseph and Carl Shapiro, "How Strong Are Weak Patents?" American Economic Review, 2008, forthcoming.

Federal Trade Commission Report, "To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy," 2003. http://www.ftc.gov/opa/2003/10/cpreport.htm.

Graham, Stuart J. H., Bronwyn H. Hall, Dietmar Harhoff and David C. Mowery, "Post-issue Patent 'Quality Control': A Comparative Study of US Patent Re-examinations and European Patent Oppositions," NBER Working Paper (8807), 2002.

Graham, Stuart J. H. and Dietmar Harhoff, "Can Post-Grant Reviews Improve Patent System Design? A Twin Study of US and European Patents," CEPR Working Paper, No. 5680, 2006.

Hall, Bronwyn H., Adam J. Jaffe and Manuel Trajtenberg, "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," NBER Working Paper, No. 8498, 2001.

Harhoff, Dietmar, Frederic Scherer and Katrin Vopel, "Citations, Family Size, Opposition and the Value of Patent Rights," Research Policy, 32, 1343-1363, 2003.

Harhoff, Dietmar and Markus Reitzig, "Determinants of Opposition Against EPO Patent Grants- the Case of Biotechnology and Pharmaceuticals," International Journal of Industrial Organization, 22, 443-480, 2004.

Howlett, Melanie J. and Andrew F. Christie, "An Analysis of the Approach of the European, Japanese and United States Patent Offices to Patenting Partial DNA Sequences (ESTs)," International Review of Industrial Property & Copyright Law, 34, 581, 2003.

Jaffe, Adam and Josh Lerner, "Innovation and its discontents : how our broken patent system is endangering innovation and progress, and what to do about it," Princeton University Press, 2004.

Jensen, Paul H., Alfons Palangkaraya and Elizabeth Webster, "Disharmony in International Patent Office Decisions," Federal Circuit Bar Journal, 2006.

Lanjouw, Jean, "Patent Protection in the Shadow of Infringement: Simulation Estimation of Patent Value," Review of Economic Studies, 65, 671-710, 1998.

Lanjouw, Jean, Ariel Pakes and Jonathan Putnam, "How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data," Journal of Industrial Economics, XLVI(4), 405-433, 1998.

Lemley, Mark , "Rational Ignorance at the Patent Office," Northwestern University Law Review, 95, 1495, 2001.

Lemley, Mark and Kimberly A. Moore, "Ending Abuse of Patent Continuations," George Mason Law & Economics Research Paper No. 03-52; UC Berkeley School of Law. Public Law Research Paper No. 140 (2003).

Lemley, Mark and Bhaven N. Sampat, "Is the Patent Office a Rubber Stamp?" Stanford Public Law Working Paper, No. 999098, 2007.

Lemley, Mark and Carl Shapiro, "Probabilistic Patents," Journal of Economic Perspectives, 19(2), 75-98, 2005.

Merges, Robert J., "As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform", Berkeley Tech. Law Journal, 14, 577-591, 1999.

Moore, Kimberly A., "Xenophobia in American Courts," Northwestern University Law Review, 97, 1497, 2003.

National Academies of Science, "A Patent System for the 21st Century," Stephen A. Merill et al. eds., 2004. http://www.nap.edu/catalog/10976.html.

Office of Inspector General, US Department of Commerce, "USPTO Should Reassess How Examiner Goals, Performance Appraisal Plans, and The Award System Stimulate and Reward Examiner

Production," 2004.

Quillen, Cecil D. and Ogden H. Webster, "Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office," Federal Circuit Bar Jouranl.,11, 1-21,2001.

Sakakibara, Mariko and Lee Branstetter, "Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms," Rand Journal of Economics, 32, 77-100, 2001.

Sampat, Bhaven, "Determinants of Patent Quality: An Empirical Analysis," Columbia University mimeo, 2005.

Shapiro, Carl, "Patent System Reform: Economic Analysis and Critique," Berkeley Technology Law Journal, 2004.

Figure 1: Time Lines for Patent Application Processes at USPTO and EPO

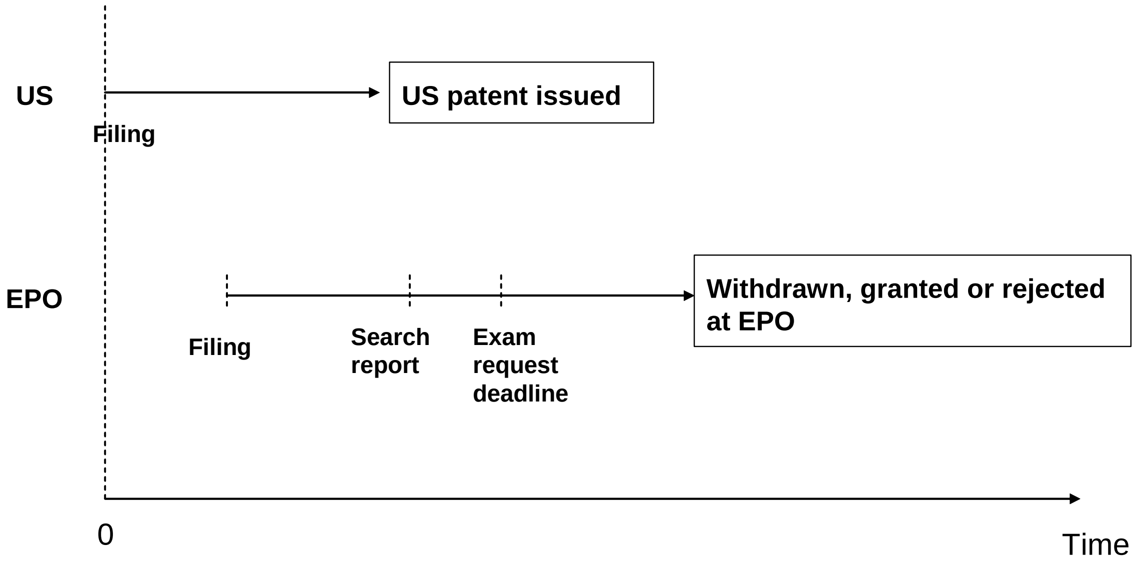


Figure 2: Decision Tree at EPO

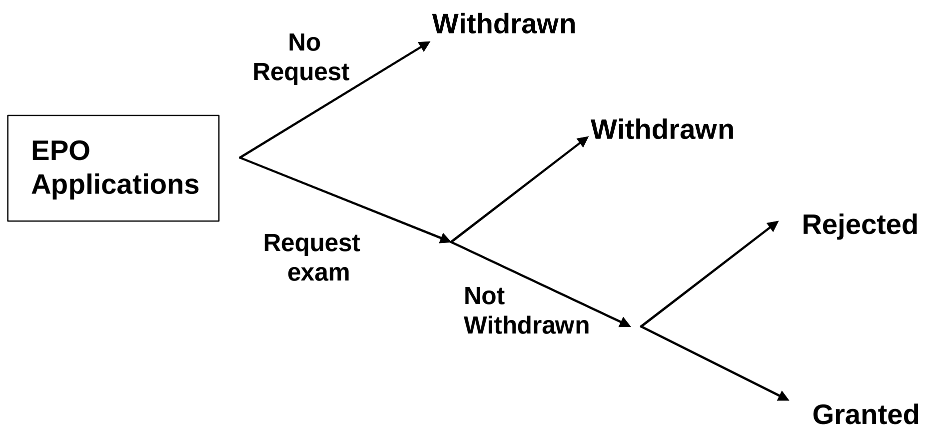
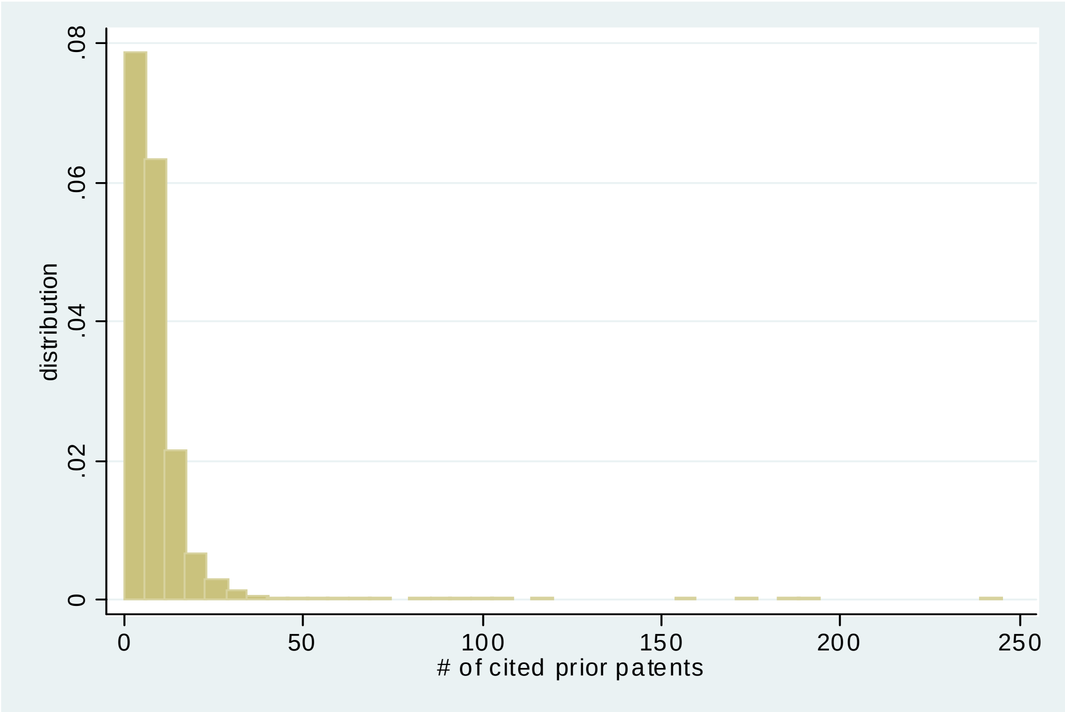
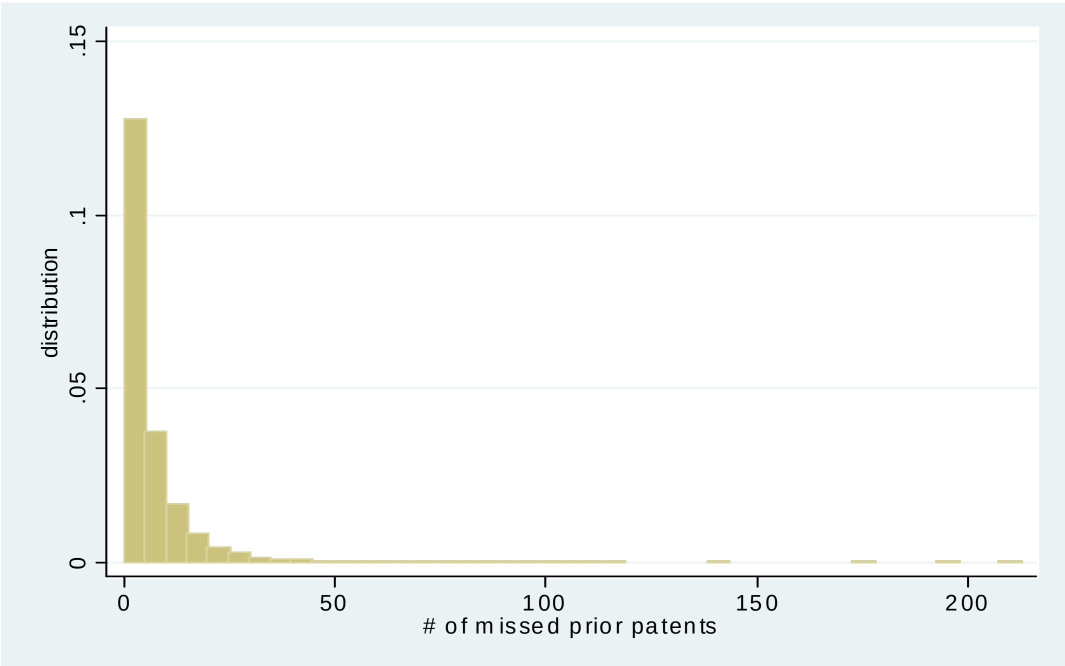


Figure 3: Distributions of cited prior patents and missed prior patents

Panel A: Histogram of the number of cited prior patents

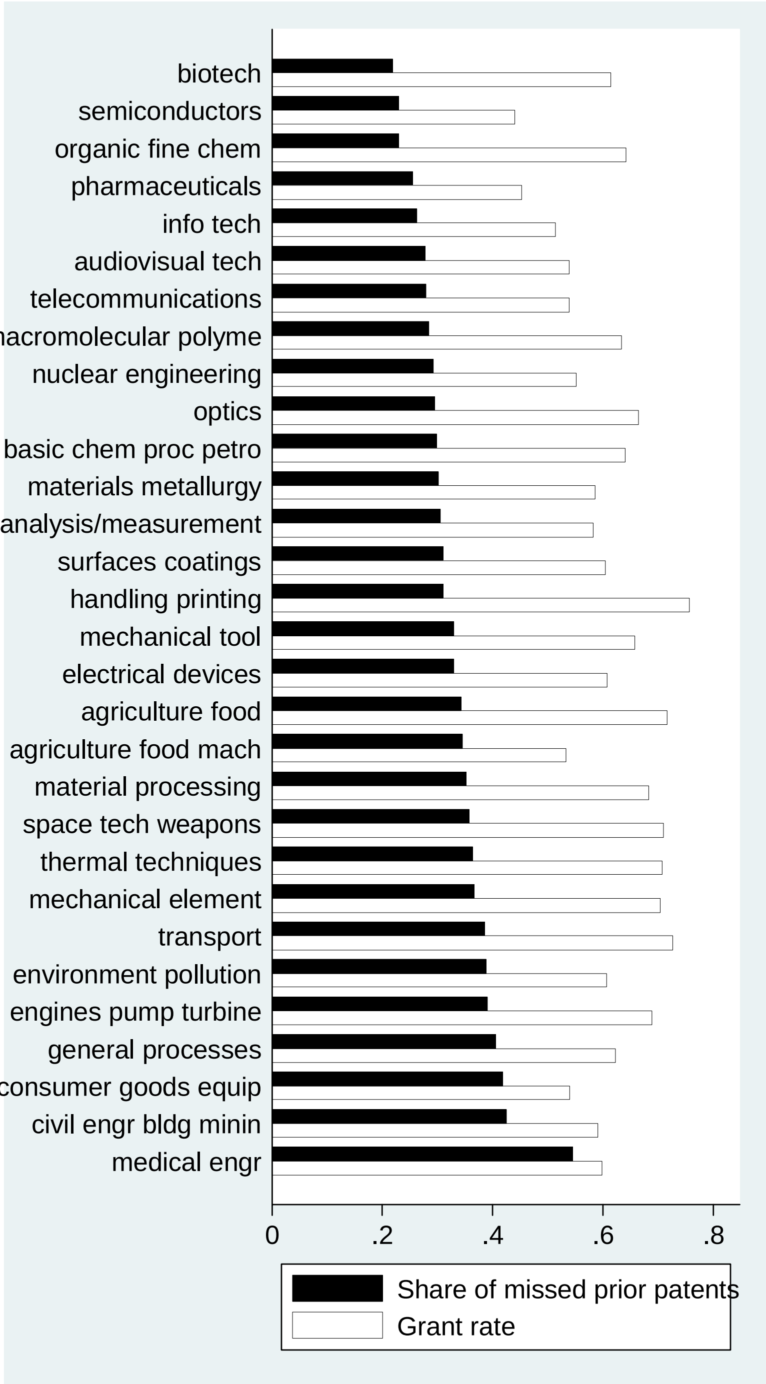


Panel B: Histogram of the number of M-CAM missed prior patents



Note: The distributions of both the number of cited prior patents and the number of M-CAM missed prior patents are concentrated in the range of 1-50 and show significant levels of dispersion, suggesting that the M-CAM algorithm has certain "intelligence" and did not mechanically retrieve a similar number of prior patents for each of US patents in the sample.

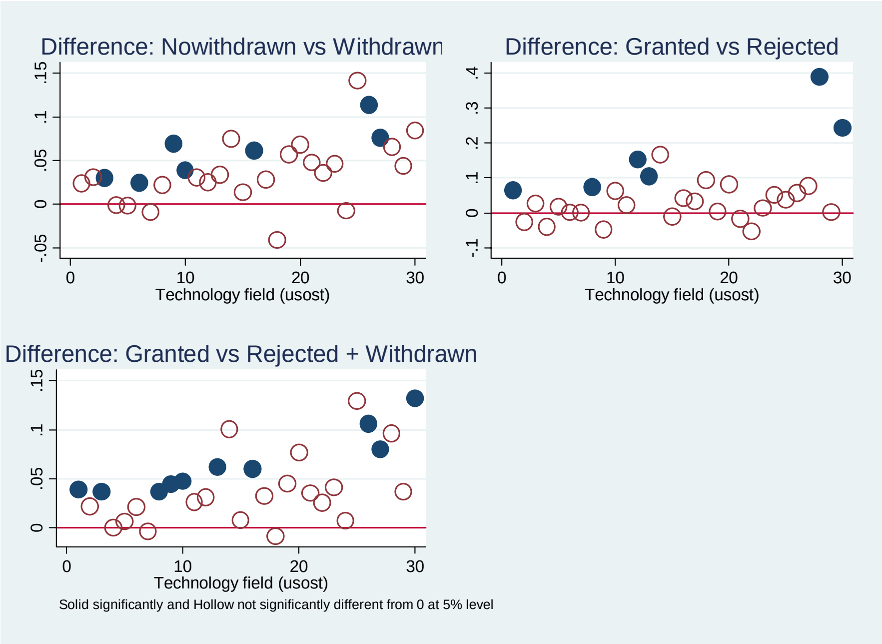
Figure 4: Comparison of average SMPP and EPO grant rates across technology fields



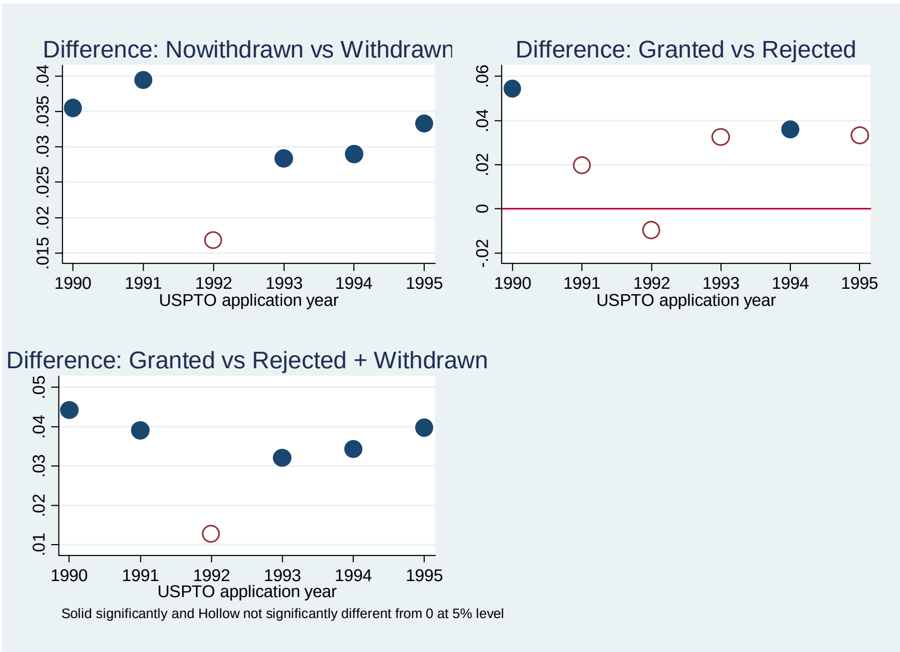
Note: The average share of missed prior patents (SMPP) and the EPO grant rates vary across technology fields. The black bars show the average SMPP across technology fields, in an ascending order. The white bars show the EPO grant rates.

Figure 5: Differences in average SMPP for US patents with different EPO application outcomes

Panel A: Differences in mean SMPP across technology fields



Panel B: Differences in mean SMPP across USPTO application years



Note: Differences in average share of missed prior patents (SMPP) for US patents with different EPO application outcomes: (1) not withdrawn versus withdrawn, (2) granted versus rejected, and (3) success (granted) versus failure (rejected and withdrawn), respectively. Panel A shows differences in average SMPP for each of the 30 technology fields and Panel B for each of the 6 USPTO application years. A solid symbol indicates a significant difference and a hollow symbol not significant, at 5% level.

Table 1: EPO application outcomes

		# of Obs	Percentage of Obs with different EPO Application Outcomes			
			withdrawn	rejected	granted	pending
(1) All inventions						
		22,317	28.32	5.57	60.24	5.87
(2) Inventions in different technology fields (OST)						
1	electrical devices	1,659	28.63	6.63	60.88	3.86
2	audiovisual tech	587	32.88	7.67	54.00	5.45
3	telecommunications	2,703	31.08	4.40	53.98	10.54
4	info tech	2,374	34.50	3.96	51.35	10.19
5	semiconductors	718	40.67	7.10	44.01	8.22
6	optics	1,961	23.05	4.95	66.50	5.51
7	analysis/measurement	1,556	29.88	5.33	58.29	6.49
8	medical engr	844	29.27	5.33	59.83	5.57
9	organic fine chem	806	24.07	7.94	64.27	3.72
10	macromolecular polyme	1,013	25.07	5.53	63.38	6.02
11	pharmaceuticals	463	36.50	11.02	45.36	7.13
12	biotech	171	28.07	4.68	61.40	5.85
13	materials metallurgy	401	29.18	7.98	58.60	4.24
14	agriculture food	74	21.62	5.41	71.62	1.35
15	general processes	747	27.31	6.69	62.38	3.61
16	surfaces coatings	754	30.11	5.84	60.48	3.58
17	material processing	467	22.48	5.57	68.31	3.64
18	thermal techniques	250	20.40	4.40	70.80	4.40
19	basic chem proc petro	337	26.11	5.64	64.09	4.15
20	environment pollution	125	31.20	6.40	60.80	1.60
21	mechanical tool	446	26.91	6.05	65.92	1.12
22	engines pump turbine	520	23.85	3.65	68.85	3.65
23	mechanical element	663	22.17	5.58	70.44	1.81
24	handling printing	1,125	17.96	4.00	75.64	2.40
25	agriculture food mach	45	31.11	11.11	53.33	4.44
26	transport	469	20.90	4.69	72.71	1.71
27	nuclear engineering	268	35.45	5.60	55.22	3.73
28	space tech weapons	90	22.22	1.11	71.11	5.56
29	consumer goods equip	427	32.08	8.90	54.10	4.92
30	civil engr bldg minin	125	30.40	6.40	59.20	4.00
99	misc unclassfied	113	23.01	5.31	67.26	4.42
(3) Inventions applied at USPTO in different years						
	1990	4,704	29.91	7.27	59.48	3.34
	1991	4,039	29.39	5.57	62.54	2.5
	1992	3,902	26.76	5.18	65.53	2.54
	1993	3,361	26.45	5.83	63.05	4.67
	1994	3,253	28.71	4.89	57.58	8.82
	1995	3,042	28.11	3.81	51.45	16.63

Note: Percentages of US patents that were withdrawn, rejected, granted and pending at the EPO, respectively. Part (1) shows the statistics is for the whole sample, Part (2) for each of the 30 technology fields, and Part (3) for each of the 6 USPTO application years.

Table 2: How good is the M-CAM NLP analysis

	Patents invalidated (US patent Number)	Prior patent that invalidates (US patent number)	Whether the invalidating prior patent is included in LLPP?	Whether the invalidating prior patent is included in MPP?
WARF Stem cell patent	5843780	5166065	Y	Y
Pfizer Lipitor patent	5969156	5273995	Y	N
Forgent JPEG	4698672	4541012	Y	Y
EpicRealm Website	5894554	5701451	Y	Y
Monsanto	5352605	4407956	Y	N
Patriot Scientific Microprocessor	5809336	4691124	N	N

Note: Patents that are revoked by the USPTO due to patentability challenges from PubPat, and the invalidating prior patents that are used by the PubPat in these challenges. In three out of six cases, the invalidating prior patents are included in "M-CAM missed prior patents" (MPP), and in five out of six, the invalidating prior patents are included in "M-CAM linguistically linked prior patents" (LLPP).

Table 3: Summary Statistics

	Panel A: The whole sample			Panel B: Analytic sample 1 in Examiner fixed effects model (examiner patent counts >=10)			Panel C: Analytic sample 2 in Examiner-tech-year fixed effects model (examiner-tech-year patent counts >=4)		
Variables	Sample size	Mean	Std. Dev.	Sample size	Mean	Std. Dev.	Sample size	Mean	Std. Dev.
SMPP	21823	0.315	0.249	18602	0.317	0.249	13291	0.319	0.249
# of cited prior patents	22300	7.709	7.260	18996	7.624	7.139	13548	7.623	6.971
# of missed prior patents	22300	5.692	9.111	18996	5.677	8.957	13548	5.802	9.183
# of total prior patents (cited and missed)	22300	13.400	14.270	18996	13.301	14.015	13548	13.425	14.073
secondexaminer_dummy	22300	0.498	0.500	18996	0.548	0.498	13548	0.577	0.494
# of linguistically linked prior patents (log)	22253	4.327	0.891	18969	4.339	0.885	13532	4.363	0.874
# of claims	21810	15.941	10.773	18786	15.957	10.833	13417	15.904	10.531
# of classifications	22300	3.957	2.711	18996	3.888	2.593	13548	3.796	2.526
# of inventors	22300	2.243	1.396	18996	2.239	1.393	13548	2.263	1.422
# of assignees	22300	0.990	0.179	18996	0.991	0.175	13548	0.993	0.164
# of total subsequent patents (citing and unciting)	22300	31.066	51.218	18996	31.558	51.560	13548	33.730	54.068
# of citing subsequent patents	22300	15.150	20.543	18996	15.597	21.060	13548	16.647	22.073
Innovation stage (total priors/total subsequents)	22199	0.351	0.242	18923	0.348	0.241	13499	0.339	0.237
Innovation stage (cited priors/citing subsequents)	22196	0.397	0.262	18922	0.392	0.260	13498	0.383	0.256
Lag to total prior patents	21823	7.182	2.819	18602	7.023	2.737	13291	6.820	2.665
Lag to cited prior patents	21745	6.479	2.985	18538	6.338	2.906	13249	6.130	2.814
Lag of total subsequent patents	21215	4.880	1.595	18116	4.956	1.536	12979	4.950	1.507
Lag of citing subsequent patents	21108	5.411	1.446	18024	5.423	1.443	12915	5.387	1.420
# of primary classes in total prior patents	22300	3.342	2.370	18996	3.318	2.344	13548	3.235	2.279
# of primary classes in cited prior patents	22300	2.719	1.895	18996	2.695	1.873	13548	2.634	1.826
# of primary classes in total subsequent patents	22300	4.400	3.721	18996	4.427	3.742	13548	4.464	3.724
# of primary classes in citing subsequent patents	22300	3.405	2.676	18996	3.441	2.698	13548	3.488	2.706

Note: Summary statistics for the full sample and the two analytical samples. The first analytical sample, used in the *US examiner* fixed effect model, consists of US patents whose examiner has at least 10 patents (examiner patent counts >=10); the second analytical sample, used in the *US examiner by technology by year* fixed effect model, consists of US patents whose examiner has at least 4 patents that are in the same technology and with the same USPTO application year (examiner-tech-year patent counts >=4).

Table 4: EPO application outcomes, using an *examiner* fixed effect model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Not Withdrawn			**Panel B: Granted, (given nonwithdrawal)**			**Panel C: Success at EPO**		
SMPP	0.1003	0.0997	0.0819	0.0454	0.0395	0.0308	0.1259	0.1220	0.0974
	(0.0171)***	(0.0178)***	(0.0185)***	(0.0129)***	(0.0135)***	(0.0143)**	(0.0182)***	(0.0195)***	(0.0204)***
Have assistant examiner	0.0045	0.0054	0.0057	0.0090	0.0097	0.0134	0.0062	0.0072	0.0090
	(0.0128)	(0.0129)	(0.0130)	(0.0091)	(0.0091)	(0.0093)	(0.0140)	(0.0140)	(0.0141)
# of LLPP (log)	-0.0011	0.0044	0.0024	-0.0030	-0.0051	-0.0024	-0.0017	0.0024	0.0021
	(0.0055)	(0.0065)	(0.0067)	(0.0042)	(0.0049)	(0.0053)	(0.0059)	(0.0068)	(0.0072)
# of claims	0.0004	0.0004	0.0004	-0.0005	-0.0005	-0.0004	-0.0002	-0.0002	-0.0001
	(0.0004)	(0.0004)	(0.0004)	(0.0003)*	(0.0003)*	(0.0003)	(0.0004)	(0.0004)	(0.0004)
# of classifications	-0.0001	-0.0004	-0.0005	-0.0008	-0.0009	-0.0006	-0.0005	-0.0009	-0.0007
	(0.0018)	(0.0018)	(0.0019)	(0.0014)	(0.0014)	(0.0013)	(0.0019)	(0.0019)	(0.0019)
# of inventors	0.0010	0.0006	0.0003	-0.0022	-0.0026	-0.0026	-0.0003	-0.0008	-0.0011
	(0.0026)	(0.0026)	(0.0026)	(0.0019)	(0.0019)	(0.0019)	(0.0029)	(0.0029)	(0.0030)
# of assignees	0.0684	0.0705	0.0730	0.0226	0.0233	0.0261	0.0766	0.0786	0.0823
	(0.0208)***	(0.0210)***	(0.0212)***	(0.0186)	(0.0186)	(0.0187)	(0.0240)***	(0.0241)***	(0.0244)***
# of total prior patents		-0.0013	-0.0004		-0.0002	-0.0000		-0.0013	-0.0003
		(0.0004)***	(0.0004)		(0.0003)	(0.0004)		(0.0004)***	(0.0005)
# of total subseq patents		0.0003	0.0001		0.0002	0.0002		0.0004	0.0002
		(0.0001)***	(0.0001)		(0.0001)***	(0.0001)***		(0.0001)***	(0.0001)
Innovation stage			-0.0634			-0.0166			-0.0709
			(0.0243)***			(0.0180)			(0.0276)**
Lag to total prior patents			-0.0011			0.0007			0.0000
			(0.0016)			(0.0012)			(0.0017)
Lag of total subseq patents			0.0021			0.0015			0.0031
			(0.0025)			(0.0020)			(0.0028)
# of primary classes in			-0.0039			-0.0017			-0.0055
total prior patents			(0.0024)			(0.0016)			(0.0026)**
# of primary classes in			0.0014			-0.0018			0.0004
total subseq patents			(0.0017)			(0.0014)			(0.0019)
Observations	18320	18320	17525	12162	12162	11664	17329	17329	16568

Note: Standard errors clustered by US examiner in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Different specifications are tested, all of which involve an examiner fixed effect. In columns 1-3, the dependent variable is 1 if an EPO application is not withdrawn and 0 if withdrawn by the applicant. In columns 4-6, the dependent variable is 1 if an EPO application is granted and 0 if rejected, conditional on it is not withdrawn. In column 7-9, the dependent variable is 1 if an EPO application succeeds (granted) at the EPO and 0 if failed (either withdrawn or rejected). Dummies for each of the 30 technology fields and for each of the 6 USPTO application years are included in all the specifications.

Table 5: EPO application outcomes, using an *examiner by technology by year* fixed effect model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Not Withdrawn			**Panel B: Granted, (given nonwithdrawal)**			**Panel C: Success at EPO**		
SMPP	0.0893	0.0902	0.0723	0.0536	0.0464	0.0381	0.1203	0.1192	0.0958
	(0.0221)***	(0.0232)***	(0.0248)***	(0.0178)***	(0.0190)**	(0.0201)*	(0.0243)***	(0.0258)***	(0.0278)***
Have assistant examiner	0.0171	0.0179	0.0184	-0.0025	-0.0015	0.0011	0.0126	0.0140	0.0149
	(0.0199)	(0.0199)	(0.0204)	(0.0159)	(0.0159)	(0.0159)	(0.0214)	(0.0214)	(0.0218)
# of LLPP (log)	-0.0030	0.0042	0.0014	-0.0078	-0.0103	-0.0085	-0.0082	-0.0016	-0.0031
	(0.0069)	(0.0081)	(0.0090)	(0.0060)	(0.0070)	(0.0077)	(0.0077)	(0.0090)	(0.0100)
# of claims	0.0004	0.0005	0.0006	-0.0008	-0.0008	-0.0006	-0.0004	-0.0003	-0.0001
	(0.0005)	(0.0005)	(0.0005)	(0.0004)*	(0.0004)**	(0.0004)	(0.0005)	(0.0005)	(0.0005)
# of classifications	0.0001	-0.0004	-0.0002	-0.0000	-0.0002	0.0005	0.0002	-0.0003	0.0004
	(0.0023)	(0.0023)	(0.0024)	(0.0020)	(0.0020)	(0.0021)	(0.0025)	(0.0025)	(0.0026)
# of inventors	-0.0000	-0.0004	-0.0008	0.0001	-0.0003	-0.0008	0.0010	0.0005	-0.0001
	(0.0034)	(0.0034)	(0.0034)	(0.0025)	(0.0025)	(0.0025)	(0.0037)	(0.0037)	(0.0038)
# of assignees	0.0711	0.0745	0.0850	0.0270	0.0279	0.0339	0.0801	0.0837	0.0955
	(0.0289)**	(0.0290)**	(0.0290)***	(0.0267)	(0.0266)	(0.0270)	(0.0329)**	(0.0331)**	(0.0333)***
# of total prior patents		-0.0017	-0.0004		-0.0003	-0.0003		-0.0018	-0.0005
		(0.0005)***	(0.0005)		(0.0005)	(0.0006)		(0.0005)***	(0.0006)
# of total subseq patents		0.0004	0.0001		0.0002	0.0003		0.0004	0.0002
		(0.0001)***	(0.0001)		(0.0001)***	(0.0001)***		(0.0001)***	(0.0002)
Innovation stage			-0.0868			-0.0093			-0.0785
			(0.0312)***			(0.0247)			(0.0348)**
Lag to total prior patents			-0.0009			0.0014			0.0001
			(0.0022)			(0.0017)			(0.0024)
Lag of total subseq patents			0.0050			0.0024			0.0067
			(0.0036)			(0.0028)			(0.0038)*
# of primary classes in total prior patents			-0.0074			-0.0004			-0.0087
			(0.0030)**			(0.0025)			(0.0034)***
# of primary classes in total subseq patents			0.0025			-0.0023			0.0018
			(0.0022)			(0.0017)			(0.0024)
Observations	13096	13096	12581	8639	8639	8308	12327	12327	11836

Note: Standard errors clustered by examiner x technology x year in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Different specifications are tested, all of which involve a *US examiner by technology by year* fixed effect. In columns 1-3, the dependent variable is 1 if an EPO application is not withdrawn and 0 if withdrawn by the applicant. In columns 4-6, the dependent variable is 1 if an EPO application is granted and 0 if rejected, conditional on it is not withdrawn. In column 7-9, the dependent variable is 1 if an EPO application succeeds (granted) at the EPO and 0 if failed (either withdrawn or rejected).

Table 6: Request for examination after search report at EPO

	(1)	(2)	(3)	(4)
	Panel A: Examiner fixed effect		**Panel B: Examiner by tech by year fixed effect**	
	Whether to request exam	Lag of request (years) (for those that requested exams)	Whether to request exam	Lag of request (years) (for those that requested exams)
SMPP	0.0252	-0.0376	0.0234	-0.0385
	(0.0104)**	(0.0168)**	(0.0145)	(0.0207)*
Have assistant examiner	0.0091	-0.0137	0.0129	-0.0196
	(0.0076)	(0.0105)	(0.0113)	(0.0173)
# of LLPP (log)	0.0026	0.0068	0.0030	0.0034
	(0.0039)	(0.0042)	(0.0051)	(0.0067)
# of claims	0.0001	-0.0002	0.0001	-0.0000
	(0.0002)	(0.0002)	(0.0002)	(0.0003)
# of classifications	-0.0000	0.0027	-0.0000	0.0035
	(0.0010)	(0.0014)**	(0.0012)	(0.0021)*
# of inventors	0.0017	0.0015	0.0034	0.0001
	(0.0013)	(0.0021)	(0.0018)*	(0.0018)
# of assignees	0.0150	-0.0261	0.0145	-0.0348
	(0.0116)	(0.0094)***	(0.0169)	(0.0143)**
# of total prior patents	-0.0000	-0.0001	0.0001	-0.0002
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
# of total subseq patents	-0.0001	0.0000	-0.0000	0.0000
	(0.0001)	(0.0000)	(0.0001)	(0.0001)
Innovation stage	-0.0216	-0.0112	-0.0147	-0.0234
	(0.0136)	(0.0185)	(0.0187)	(0.0308)
Lag to total prior patents	-0.0008	0.0010	-0.0001	0.0013
	(0.0009)	(0.0016)	(0.0012)	(0.0025)
Lag of total subseq patents	0.0008	0.0016	0.0015	0.0032
	(0.0015)	(0.0012)	(0.0021)	(0.0020)
# of primary classes in	-0.0016	-0.0003	-0.0027	0.0019
total prior patents	(0.0013)	(0.0016)	(0.0017)	(0.0024)
# of primary classes in	0.0002	-0.0012	0.0011	-0.0013
total subseq patents	(0.0009)	(0.0012)	(0.0013)	(0.0012)
Observations	15198	14325	10882	10279

Note: Standard errors, clustered by examiner in Panel A and by examiner x technology x year in Panel B, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Panel A involves an *US examiner* fixed effect model. The dependent variable in Column 1is 1 if an applicant requests examination after seeing the EPO search report and 0 if he does not. The dependent variable in Column 2 is the lag of exam request, measured by the period between search report publication date and exam request date. The dependent variables in Columns 3-5 are EPO outcomes, conditional on a request for examination being made after the search report. Panel B involves a *US examiner by technology by year* fixed effect model.

Table 7: Testing the first alternative rationales of information flow

EPO outcomes for observations with search report PRIOR to US patent issuance

	(1)	(2)	(3)	(4)	(5)	(6)
	(a) examiner fixed effect			(b) examiner-tech-year fixed effect		
	Not withdrawn	Granted, given nowithdrawal	Success at EPO	Not withdrawn	Granted, given nowithdrawal	Success at EPO
SMPP	0.1277	0.0506	0.1451	0.1556	0.0507	0.1772
	(0.0372)***	(0.0261)*	(0.0405)***	(0.0566)***	(0.0400)	(0.0624)***
Have assistant examiner	0.0101	0.0187	0.0282	-0.0114	0.0400	0.0186
	(0.0255)	(0.0194)	(0.0263)	(0.0453)	(0.0458)	(0.0511)
# of LLPP (log)	-0.0071	-0.0050	-0.0075	-0.0149	-0.0153	-0.0197
	(0.0143)	(0.0099)	(0.0149)	(0.0214)	(0.0170)	(0.0235)
# of claims	0.0006	-0.0010	-0.0002	0.0006	-0.0005	-0.0000
	(0.0006)	(0.0005)**	(0.0006)	(0.0010)	(0.0008)	(0.0010)
# of classifications	0.0023	0.0028	0.0035	0.0009	0.0031	0.0020
	(0.0033)	(0.0024)	(0.0036)	(0.0051)	(0.0045)	(0.0058)
# of inventors	-0.0017	-0.0013	-0.0019	0.0034	-0.0020	0.0049
	(0.0045)	(0.0036)	(0.0055)	(0.0062)	(0.0050)	(0.0069)
# of assignees	0.0549	0.0456	0.0814	0.0824	0.0815	0.1171
	(0.0377)	(0.0306)	(0.0430)*	(0.0603)	(0.0502)	(0.0698)*
# of total prior patents	-0.0013	-0.0008	-0.0015	-0.0008	-0.0009	-0.0010
	(0.0008)	(0.0007)	(0.0008)*	(0.0010)	(0.0012)	(0.0012)
# of total subseq patents	0.0001	0.0003	0.0001	-0.0000	0.0003	-0.0000
	(0.0002)	(0.0001)*	(0.0002)	(0.0003)	(0.0002)	(0.0003)
Innovation stage	-0.0597	-0.0236	-0.0795	-0.1223	-0.0093	-0.1429
	(0.0497)	(0.0339)	(0.0535)	(0.0638)*	(0.0514)	(0.0721)**
Lag to total prior patents	-0.0040	0.0002	-0.0033	-0.0002	0.0012	0.0012
	(0.0029)	(0.0023)	(0.0032)	(0.0047)	(0.0035)	(0.0051)
Lag of total subseq patents	-0.0027	0.0026	0.0007	0.0000	0.0027	0.0034
	(0.0051)	(0.0036)	(0.0054)	(0.0080)	(0.0061)	(0.0083)
# of primary classes in	-0.0024	0.0016	-0.0013	-0.0035	0.0013	-0.0029
total prior patents	(0.0044)	(0.0029)	(0.0047)	(0.0060)	(0.0047)	(0.0067)
# of primary classes in	0.0028	-0.0028	0.0015	0.0031	-0.0032	0.0012
total subseq patents	(0.0033)	(0.0023)	(0.0038)	(0.0044)	(0.0038)	(0.0053)
Observations	5153	3454	4917	3117	2075	2954

Note: Standard errors, clustered by the fixed effects, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Columns 1-6 test the alternative rationale of information flow from the USPTO to the EPO, by studying observations whose EPO search reports were prior to the issuance of corresponding US patents. Columns 1-3 involve an examiner fixed effect with dummies for technology fields and USPTO application years; Columns 4-6 involve an examiner-technology-year fixed effect. The dependent variables are the same as in Tables 4 and 5. The dependant variables are in years.

Table 8: Testing the second alternative rationale of difficulty in finding prior art

EPO Search Interval (years)

	(1)	(2)	(3)	(4)
	(a) examiner fixed effect		(b) examiner-tech-year fixed effect	
	Total search interval	Search interval per citation	Total search interval	Search interval per citation
SMPP	-0.0812	-0.0286	-0.0830	-0.0276
	(0.0271)***	(0.0153)*	(0.0369)**	(0.0222)
Have assistant examiner	0.0035	0.0006	0.0002	0.0053
	(0.0225)	(0.0134)	(0.0320)	(0.0195)
# of LLPP (log)	0.0260	-0.0165	0.0293	-0.0063
	(0.0125)**	(0.0059)***	(0.0155)*	(0.0078)
# of claims	0.0016	0.0004	0.0033	0.0006
	(0.0006)***	(0.0003)	(0.0008)***	(0.0004)
# of classifications	0.0009	-0.0014	0.0041	-0.0007
	(0.0027)	(0.0014)	(0.0033)	(0.0019)
# of inventors	0.0085	0.0006	0.0123	0.0022
	(0.0048)*	(0.0027)	(0.0063)*	(0.0038)
# of assignees	-0.0445	0.0062	-0.0669	-0.0102
	(0.0307)	(0.0150)	(0.0431)	(0.0198)
# of total prior patents	-0.0009	-0.0004	-0.0012	-0.0005
	(0.0007)	(0.0003)	(0.0009)	(0.0004)
# of total subseq patents	0.0002	0.0002	0.0004	0.0002
	(0.0002)	(0.0001)	(0.0003)	(0.0002)
Innovation stage	-0.0317	-0.0466	-0.0687	-0.0555
	(0.0404)	(0.0207)**	(0.0507)	(0.0269)**
Lag to total prior patents	-0.0079	-0.0070	-0.0070	-0.0067
	(0.0029)***	(0.0015)***	(0.0037)*	(0.0018)***
Lag of total subseq patents	-0.0033	-0.0014	-0.0055	-0.0029
	(0.0046)	(0.0020)	(0.0059)	(0.0027)
# of primary classes in	-0.0055	-0.0060	-0.0031	-0.0086
total prior patents	(0.0043)	(0.0019)***	(0.0052)	(0.0027)***
# of primary classes in	0.0068	0.0024	0.0001	0.0010
total subseq patents	(0.0036)*	(0.0018)	(0.0043)	(0.0023)
Observations	17422	16725	12495	11948

Note: Standard errors, clustered by the fixed effects, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Columns 1-4 test the alternative rationale of difficulty of finding prior art, by studying the total search interval and the search interval per citation at the EPO search office. Columns 1-2 involve an examiner fixed effect with dummies for technology fields and USPTO application years; Columns 3-4 involve an examiner-technology-year fixed effect. The dependant variables are in years.

Table 9: Testing the third alternative: Effects of cited prior patents and missed prior patents on EPO application outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Not withdrawn at EPO			**Panel B: Granted, given non-withdrawal**			**Panel C: Success at EPO**		
(a) Examiner fixed effect									
# of cited prior patents	-0.0018		-0.0022	-0.0011		-0.0015	-0.0023		-0.0030
	(0.0008)**		(0.0008)***	(0.0007)		(0.0007)**	(0.0009)***		(0.0009)***
# of missed prior patents		0.0011	0.0015		0.0009	0.0012		0.0018	0.0024
		(0.0005)**	(0.0005)***		(0.0004)**	(0.0004)***		(0.0006)***	(0.0006)***
control variables	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	17525	17525	17525	11664	11664	11664	16568	16568	16568
(b) Examiner by technology by year fixed effect									
# of cited prior patents	-0.0019		-0.0023	-0.0015		-0.0019	-0.0028		-0.0034
	(0.0010)*		(0.0010)**	(0.0011)		(0.0011)*	(0.0012)**		(0.0012)***
# of missed prior patents		0.0010	0.0013		0.0008	0.0011		0.0016	0.0022
		(0.0007)	(0.0008)*		(0.0006)	(0.0007)*		(0.0008)**	(0.0009)**
control variables	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	12581	12581	12581	8308	8308	8308	11836	11836	11836

Note: Standard errors, clustered by the fixed effects, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Test the effects of the number of cited prior art and the number of missed prior patents on EPO application outcomes. Part (a) involves a *US examiner* fixed effect model and Part (b) a *US examiner by technology by year* fixed effect. The dependent variables in Columns 1-3 are 1 if not withdrawn at the EPO and 0 otherwise. The dependent variables in Columns 4-6 are 1 if granted and 0 if rejected, conditional on non-withdrawal. The dependent variables in Columns 7-9 are 1 if success (granted) at the EPO and 0 if failed (withdrawn or rejected). Control variables are the same controls in Table 4 and 5, with the number of total prior patents excluded. Dummies for each of the 30 technology fields and for each of the 6 USPTO application years are included in Part (a) that involve a *US examiner* fixed effect.

Table 10: EPO application outcomes, using a *US examiner by patent assignee* fixed effect model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Not Withdrawn			**Panel B: Granted, (given nonwithdrawal)**			**Panel C: Success at EPO**		
SMPP	0.0698	0.0638	0.0517	0.0633	0.0563	0.0480	0.1095	0.0994	0.0766
	(0.0287)**	(0.0295)**	(0.0310)*	(0.0200)***	(0.0211)***	(0.0225)**	(0.0315)***	(0.0327)***	(0.0344)**
Have assistant examiner	0.0026	0.0037	0.0040	0.0132	0.0147	0.0144	-0.0022	-0.0007	0.0007
	(0.0231)	(0.0231)	(0.0232)	(0.0197)	(0.0197)	(0.0204)	(0.0242)	(0.0242)	(0.0245)
# of LLPP (log)	0.0030	0.0085	-0.0002	-0.0009	-0.0025	-0.0075	0.0038	0.0087	-0.0011
	(0.0089)	(0.0106)	(0.0113)	(0.0074)	(0.0087)	(0.0092)	(0.0098)	(0.0115)	(0.0121)
# of claims	0.0013	0.0013	0.0012	-0.0006	-0.0006	-0.0005	0.0006	0.0006	0.0007
	(0.0005)**	(0.0005)**	(0.0005)**	(0.0005)	(0.0005)	(0.0005)	(0.0006)	(0.0006)	(0.0006)
# of classifications	0.0020	0.0013	0.0018	0.0022	0.0019	0.0020	0.0036	0.0027	0.0029
	(0.0028)	(0.0029)	(0.0029)	(0.0023)	(0.0023)	(0.0023)	(0.0032)	(0.0032)	(0.0032)
# of inventors	0.0032	0.0029	0.0022	0.0009	0.0007	0.0003	0.0048	0.0044	0.0038
	(0.0042)	(0.0042)	(0.0043)	(0.0030)	(0.0031)	(0.0031)	(0.0043)	(0.0044)	(0.0044)
# of assignees	-0.1014	-0.1031	-0.1189	-0.0046	-0.0028	-0.0113	-0.0931	-0.0948	-0.1149
	(0.0997)	(0.0997)	(0.0959)	(0.0339)	(0.0343)	(0.0407)	(0.1030)	(0.1028)	(0.0984)
# of total prior patents		-0.0015	-0.0008		-0.0002	0.0001		-0.0016	-0.0005
		(0.0005)***	(0.0006)		(0.0005)	(0.0005)		(0.0006)***	(0.0006)
# of total subseq patents		0.0004	0.0003		0.0002	0.0002		0.0005	0.0003
		(0.0002)**	(0.0002)		(0.0001)**	(0.0001)*		(0.0002)***	(0.0002)
Innovation stage			-0.0629			0.0320			-0.0417
			(0.0392)			(0.0324)			(0.0467)
Lag to total prior patents			0.0001			0.0013			0.0017
			(0.0026)			(0.0021)			(0.0029)
Lag of total subseq patents			0.0067			0.0007			0.0065
			(0.0043)			(0.0031)			(0.0047)
# of primary classes in total prior patents			-0.0013			-0.0044			-0.0069
			(0.0038)			(0.0027)			(0.0040)*
# of primary classes in total subseq patents			0.0011			0.0015			0.0029
			(0.0032)			(0.0023)			(0.0034)
Observations	8152	8152	7830	5368	5368	5169	7649	7649	7343

Note: Standard errors clustered by examiner x assignee in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Different specifications are tested, all of which involve a *US examiner by patent assignee* fixed effect. In columns 1-3, the dependent variable is 1 if an EPO application is not withdrawn and 0 if withdrawn by the applicant. In columns 4-6, the dependent variable is 1 if an EPO application is granted and 0 if rejected, conditional on it is not withdrawn. In column 7-9, the dependent variable is 1 if an EPO application succeeds (granted) at the EPO and 0 if failed (either withdrawn or rejected). Dummies for each of the 30 technology fields and for each of the 6 USPTO application years are included.

Table 11: Do *primary examiner by secondary examiner* fixed effects make difference?

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Primary and Secondary Examiners fixed effects			**Panel B: Primary and Secondary Examiners by Technology by Year fixed effects**		
	Not withdrawn at EPO	Granted, given non-withdrawal	Success at EPO	Not withdrawn at EPO	Granted, given non-withdrawal	Success at EPO
(a) For patents either with a primary examiner or with both a primary and a secondary examiner						
SMPP	0.0710	0.0214	0.0802	0.0613	0.0394	0.0747
	(0.0210)***	(0.0173)	(0.0237)***	(0.0304)**	(0.0271)	(0.0344)***
Observations	14044	9465	13302	9584	6404	9023
(b) For patents with both a primary examiner and a secondary examiner						
SMPP	0.0875	0.0478	0.1141	0.1047	0.0359	0.1082
	(0.0366)**	(0.0297)	(0.0414)**	(0.0529)**	(0.0452)	(0.0604)*
Observations	5133	3427	4864	3486	2295	3288
Examiner fixed effect	yes	yes	yes	no	no	no
Tech and Year dummies	yes	yes	yes	no	no	no
Examiner-tech-year fixed effect	no	no	no	yes	yes	yes
Control variables	yes	yes	yes	yes	yes	yes

Note: Standard errors, clustered by the fixed effects, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Part (a) includes patents that either have a primary examiner alone or have both primary and secondary examiners, and part (b) looks at patents with both primary and secondary examiners. Panel A in Columns 1-3 involve a *US examiner* fixed effect, with dummies for technology fields and USPTO application years. Panel B in Columns 4-6 involve a *US examiner by technology by year* fixed effect.

Table 12: JPO application outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Examiner fixed effect			**Panel B: Examiner-tech-year fixed effect**		
	Not withdrawn at JPO	Granted, given nonwithdrawal	Success at JPO	Not withdrawn at JPO	Granted, given nonwithdrawal	Success at JPO
SMPP	0.0472	0.0141	0.0455	0.0515	0.0240	0.0518
	(0.0189)**	(0.0235)	(0.0218)**	(0.0243)**	(0.0345)	(0.0282)*
Have assistant examiner	-0.0074	0.0177	0.0112	0.0084	0.0423	0.0350
	(0.0141)	(0.0158)	(0.0141)	(0.0205)	(0.0272)	(0.0228)
# of LLPP (log)	0.0048	-0.0082	0.0005	0.0079	-0.0129	-0.0065
	(0.0074)	(0.0097)	(0.0086)	(0.0095)	(0.0135)	(0.0111)
# of claims	-0.0007	0.0004	-0.0006	-0.0005	-0.0001	-0.0007
	(0.0003)**	(0.0005)	(0.0004)	(0.0005)	(0.0006)	(0.0005)
# of classifications	-0.0017	-0.0004	-0.0017	-0.0025	-0.0005	-0.0035
	(0.0018)	(0.0027)	(0.0021)	(0.0024)	(0.0038)	(0.0029)
# of inventors	0.0146	0.0154	0.0241	0.0158	0.0131	0.0243
	(0.0027)***	(0.0032)***	(0.0030)***	(0.0031)***	(0.0046)***	(0.0039)***
# of assignees	0.0348	-0.0299	0.0123	0.0360	-0.0757	-0.0174
	(0.0218)	(0.0330)	(0.0275)	(0.0305)	(0.0409)*	(0.0368)
# of total prior patents	0.0001	0.0010	0.0005	0.0002	0.0015	0.0011
	(0.0005)	(0.0006)*	(0.0005)	(0.0005)	(0.0009)*	(0.0007)*
# of total subseq patents	0.0002	0.0003	0.0004	0.0001	0.0000	0.0002
	(0.0001)	(0.0001)**	(0.0001)***	(0.0001)	(0.0002)	(0.0002)
Innovation stage	-0.0796	-0.0924	-0.1060	-0.1145	-0.0763	-0.1254
	(0.0251)***	(0.0334)***	(0.0294)***	(0.0317)***	(0.0501)	(0.0389)***
Lag to total prior patents	-0.0034	0.0037	-0.0006	-0.0016	0.0019	0.0009
	(0.0016)**	(0.0022)*	(0.0019)	(0.0021)	(0.0031)	(0.0025)
Lag of total subseq patents	0.0038	0.0090	0.0083	0.0046	0.0042	0.0058
	(0.0028)	(0.0042)**	(0.0033)**	(0.0038)	(0.0055)	(0.0043)
# of primary classes in	-0.0007	-0.0005	-0.0003	-0.0011	0.0022	-0.0012
total prior patents	(0.0026)	(0.0032)	(0.0029)	(0.0032)	(0.0046)	(0.0039)
# of primary classes in	0.0019	-0.0024	-0.0008	0.0012	-0.0011	-0.0003
total subseq patents	(0.0018)	(0.0021)	(0.0020)	(0.0022)	(0.0030)	(0.0027)
Observations	17574	10208	15450	12625	7491	11155

Note: Standard errors clustered by the fixed effects in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Panel A in Columns 1-3 involve a *US examiner* fixed effect, with dummies for technology fields and USPTO application years. Panel B in Columns 4-6 involve a *US examiner by technology by year* fixed effect. The dependent variables in Columns 1 and 4 are is 1 if a JPO application is not withdrawn and 0 if withdrawn by the applicant. The dependent variables in Columns 2 and 5 are 1 if a JPO application is granted and 0 if rejected, conditional on it is not withdrawn. The dependent variables in Columns 3 and 6 are 1 if a JPO application succeeds (granted) at the JPO and 0 if failed (either withdrawn or rejected).

Appendix A: M-CAM NLP Algorithm

For a given root patent, M-CAM algorithm provides five sets of patents: (1) patents that are cited by the root patents (cited prior patents); (2) patents that cite the root patents (citing subsequent patents); (3) patents that are linguistically (and technically) related, but were not cited (uncited prior patents); (4) patents that are linguistically (and technically) related, issued later, and did not cite the root patent (unciting subsequent patents); and (5) patents that are linguistically related, and being prosecuted by the USPTO at the same time as the root patent (concurrent art).

For clarity, we refer to an example below. Figure 1 illustrates the sets of patents that M-CAM's algorithm detects for a US patent (US5319702). Across the bottom of the figure, is a time line from 1973 to 2005. The height of the "mountains" represents the relative number of patents issued in the year in a particular category. In the M-CAM interface, the "mountains" are colour coded. Here, for lack of color, they are different shades of grey.

The dark grey mountain in the middle represents "concurrent art." US5319702 was going through the patent office during the years 1992 to 1996 and the dark grey mountain represents patents that are linguistically related to US '702 according to the algorithm and for which the applications were proceeding through the patent office at the same time as the '702 patent. It is likely that examiners may therefore have missed this body of art, regardless of its relevance to the root patent.

The low, black "mountains" indicated cited and citing patents. The black area to the left of the concurrent art area indicates patents that are cited by the root patent (cited prior patents), and the black area to the right of the concurrent art area indicates patents that cite the '702 patent (citing subsequent patents). The light grey mountains to the left and right side of the concurrent art area indicate, respectively: patents that the algorithm finds linguistically related but are not cited by the root patent (uncited prior patents), and patents that the algorithm finds linguistically related but that do not cite the root patent (unciting subsequent patents).

The linguistically linked previous patents (LLPP), used in the study, are the light grey mountain to the left of the concurrent art area. Among these, M-CAM algorithm also highlights those whose similarity to the root patent is very significant. Those highlighted LLPP are used as M-CAM missed prior patents to construct the variable, share of missed prior patents (SMPP). Of course, the algorithm is not comparable to a legal or technical assessment of whether a previous patent "should have been cited." However, on average, and over a large data set, the algorithm does provide a useful tool for understanding technically related patents.

Figure A1: M-CAM Innovation Space