Centre for Transformative Innovation

Working Paper Series

INNOVATION = IDEAS + CHANGE

SWINBURNE UNIVERSITY OF TECHNOLOGY

Low-quality patents in the eye of the beholder: Evidence from multiple examiners

Working Paper 1/16

Gaétan de Rassenfosse, Adam B. Jaffe and Elizabeth Webster

January 2016

Low-quality patents in the eye of the beholder: Evidence from multiple examiners

Gaétan de Rassenfosse
École polytechnique fédérale de Lausanne (EPFL), Chair of Innovation and IP Policy,
College of Management of Technology, Odyssea 2.01.1, Station 5, 1015 Lausanne, Switzerland.
Email: gaetan.derassenfosse@epfl.ch (corresponding author).

Adam B. Jaffe
Director and Senior Fellow, Motu Economic and Public Policy Research; Adjunct Professor, Queensland University of Technology; Economic and Social Systems Research Theme Leader,
Te Punaha Matatini Centre of Research Excellence. Email: Adam.Jaffe@motu.org.nz

Elizabeth Webster
Centre for Transformative Innovation, Swinburne University of Technology,
H25, PO Box 218 Hawthorn, Victoria 3122 Australia. Email: emwebster@swin.edu.au

This version: January 2016

Abstract

A low-quality patent system threatens to slow the pace of technological progress. Concerns about low patent quality are supported by estimates from litigation studies suggesting that the majority of patents granted by the U.S. patent office should not have been issued. This paper proposes a new way of measuring patent quality, based on twin patent applications granted at one office but refused at another office, applied to the five largest patent offices. The results suggest that quality in patent systems is higher than previously thought, although the U.S. patent office's performance is poorer than those of Europe and Japan.

Alternate Abstract

A low-quality patent system threatens to slow the pace of technological progress. Concerns about low patent quality are supported by estimates from litigation studies suggesting that the majority of patents granted by the U.S. patent office should not have been issued. This paper proposes a new way of measuring patent quality, based on twin patent applications granted at one office but refused at another office, applied to the five largest patent offices. Our method allows us to distinguish low-quality patents issued because an office has a (consistent) low standard from patents issued in violation of an office's own standard, however high or low (so-called 'weak patents').The results suggest that quality in patent systems is higher than previously thought; in particular the percentage of 'weak' patents is in single digits for all offices, although the U.S. patent office's performance is poorer than those of Europe and Japan.

Keywords: inventive step, non-obviousness, patent quality, weak patent

JEL codes: O34, L43, K41

1. Introduction

Concern that the patent system inhibits rather than encourages innovation has become a staple of the business and technology press (*e.g.*, *The Economist*, 2015). A major source of concern is that patent offices may grant too many low-quality patents, whose existence can chill the R&D investment and commercialization processes, either because of background uncertainty about freedom to operate or because of implicit or explicit threats of litigation.

Concern about patent quality is by no means new. The recent *Economist* article quoted itself from 1851 saying that the granting of patents "begets disputes and quarrels betwixt inventors, provokes endless lawsuits [and] bestows rewards on the wrong persons." But in the last few decades, significant increases in the number of patent applications granted and the frequency of patent litigations, as well as media attention such cases have received, have given these concerns new force in the academic literature. Major patent offices are well aware of the problem and several of them have initiatives underway aimed at improving the quality of patent review. For example, the U.S. Patent and Trademark Office (USPTO) now has an Office of Patent Quality Assurance and has recently initiated an ongoing online 'patent quality chat.'[1]

We interpret concern about low-quality patents as corresponding to concern that patents are being granted whose inventive step is too small to deserve patent protection. Conceptually, there are two pathways by which this may be occurring. A first source of low quality in a patent system relates to the fact that patent offices might *systematically apply a standard that is too lenient*, relative to some conception of optimal stringency. Some of the discussion of the patent quality problem, particularly in the United States, has this flavor. Jaffe and Lerner (2004), for example, argue that changes in the incentives of the USPTO, the U.S. courts, and U.S. patentees over the 1980s and 1990s led to a systematic lowering of the standard for a U.S. patent grant.

A conceptually distinct source of low quality in patent system is mistakes—*granting patents that in actuality do not meet the office's own implicit standard*, however high or low that standard may be. Observers of the patent system also discuss this issue. For example, Lemley and Shapiro (2005:83) write: "There is widespread and growing concern that the Patent and Trademark Office issues far too many 'questionable' patents that are unlikely to be found valid based on a thorough review." Although there are clear patentability requirements and patentable subject matters, flaws in the examination process (Meurer, 2009; Lemley and Sampat, 2012; Frakes and Wasserman, forthcoming; Nagaoka and Yamauchi, 2015) and in the governance of patent offices (de Saint-Georges and van Pottelsberghe, 2013; Picard and van Pottelsberghe, 2014) affect the quality of the examination process. More generally, the grant decision rests ultimately on a subjective

[1] See <https://www.uspto.gov/patent/initiatives/2016-patent-quality-chats>

comparison of the application's inventive merit and the office's standard for novelty. Perfect consistency of decision-making seems unlikely to be the outcome of such a process.

The practical and normative consequences of these different sources of low quality are different. Systematically low standards create monopoly power and transfer rents in situations where the triviality of the invention arguably does not justify the reward. But low standards consistently applied are not, logically, a source of uncertainty about which patents are truly valid—so long as the patent office and the courts are applying exactly the same standard. Such uncertainty only comes about if standards are not applied consistently. Scholarly literature refers to patents that were granted because standards were not applied consistently as 'weak' patents. It argues that the litigation threat that they pose reduces welfare by leading consumers to pay supra-competitive prices due to the public good nature of challenging a patent (Farrell and Shapiro, 2008; Encaoua and Lefouili, 2009; Choi and Gerlach, 2016).

We propose a formal model that attributes inconsistent patent examination decisions across offices to systematic differences in offices' propensity to grant applications (capturing *de facto* policies and practices) or mistakes by one or another office. We then use novel data on multiple examination outcomes for the same invention in different patent offices to estimate the magnitude of these sources of inconsistency. Our data are derived from a population of about 400,000 inventions with linked patent applications that have been examined in at least two of the five major patent offices, covering in total more than a million applications. The premise of our model is that a refusal by an examiner in one jurisdiction raises doubts with regard to the legitimacy of the patent grant secured elsewhere. In particular, we estimate a statistical model of the grant process that captures parametrically the effect of observable application attributes on the grant probability, the effect of systematic differences in propensity to grant applications across offices, and the possibility of personal (*i.e.*, examiner) discretion in every decision.

To foreshadow the results, we find that systematic differences across offices appear to be larger than within-country inconsistency of decisions, but such inconsistency is present to varying degrees across countries. The model estimates imply that only 2–6 percent of granted patents have dubious validity in the specific sense that they appear to be inconsistent with the country's own standard for patent grant (what we call a weak patent). An additional 2-15 percent can be thought of as low-quality in the sense that they would not have been granted by the strictest office. Patent offices in China and the United States appear to be the most lenient offices, and the Japan patent office the strictest. While these estimates are of interest in their own rights, given the difficulty in measuring patent quality, they also inform policy discussion. In particular, our results have important implications for current international agreements between patent offices and for discussions about how to fix the patent system.

The rest of the paper is organized as follows. Section 2 presents background discussion on patent quality. Section 3 presents the empirical strategy and Section 4 presents the data. Sections 5 and 6 discuss the econometric results and robustness tests, respectively. Section 7 concludes.

2. Background

Most of the existing literature looks at the issue of low quality by measuring the fraction of litigated patents that are found by a court to be invalid. Such studies provide valuable insights on the prevalence of invalidity. It is unclear, however, how invalidation in court relates to the two possible sources of invalidity. If one assumes that the courts are implicitly applying the same standard as the patent office, and that courts make perfect decisions, then a court invalidity finding corresponds to a case in which the office did not correctly apply its own standard. In practice, it is also possible that the court is applying a more stringent standard—and that it makes mistakes (Lemley, 2001). Thus, litigation studies tell little about the quality of the examination process or the stringency of the office.

Nonetheless, patent litigation studies report 'invalidity' rates in the range between 30 to 75 percent. Allison and Lemley (1998) reviewed final validity decisions of 299 litigated patents and found an invalidity rate of half. Cremers *et al.* (2014) report that about 30 percent of appealed patent suits have their initial decision overturned. Furthermore, European patents, with the same set of claims, that are litigated in multiple courts can differ in their court outcome. Zischka and Henkel (2014) affirm this high rate of uncertainty and find a 75 percent invalidity rate of appeals at the German Federal Patent Court between 2000 and 2012. These studies suggest that invalidity rates might be quite high. However, given that a mere 0.1 percent of patents are litigated to trial (Lemley and Shapiro, 2005), such patents are not a random sample of the population, so it remains unclear what these statistics tell us about the overall prevalence of invalidity. This point is well made by Marco (2004), who emphasizes the importance of accounting for selection effects in patent validity adjudications.

Recognising this problem, Miller (2013) attempts to correct for selection into an invalidity hearing. Using 980 adjudicated and 1960 control patents at the USPTO, he estimates a population-wide invalidity rate of 28 percent. However, the selection into Miller's sample is twofold: selection into a patent being disputed, and selection into parties choosing trial over settlement. The first selection is not accounted for, suggesting that the 28-percent figure may still be biased, though the direction of bias is unclear. Zischka and Henkel (2014) have also studied the presence of selection bias in their data but did not identify statistically significant selection covariates. More recently, scholars have also studied the outcome of *inter partes* reviews, which are post-grant reviews conducted by USPTO Patent Trial and Appeal Board (Wallach and Darrow, 2016). There are also selection effects at play, which one should properly model in order to obtain population-wide estimates of invalidity.

As illustrated by the litigation studies, the basic approach to assessing the level of quality in the system is to investigate what happens when another qualified decision maker (but ideally many) takes a fresh look at the question of whether an asserted invention qualifies for patent protection. As far as we can ascertain, the only academic study in that vein that does not rely on litigation data is Paradise et al. (2005). The authors manually examine the validity of 1167 claims of 74 U.S. patents on human genetic material. They find that 448 claims (38%) were problematic. The 'second-pair-of-eyes review' program at the USPTO, which began in the year 2000 but has been discontinued since, aims at assessing examination quality by re-examining patent applications related to business methods. However, data are not publicly available and Allison and Hunter (2006:737-8) comment that this review is a "subjective, in-house process metric guided by no apparent standards that may fall victim to unconscious bias or external influence."

In contrast with these studies, Palangkaraya, Jensen and Webster (2011) use a revealed behavior method to estimate rates of patent invalidity. They analyze the population of all 34,000 patent applications that were granted by the USPTO and examined at both the EPO and JPO during the 1990s. Assuming that the number of forward citations at the USPTO is a proxy for the real size of the inventive step, they estimate that 6.1 and 9.8 percent of patents are, respectively, incorrectly rejected and incorrectly granted.

Finally, note that other studies have empirically examined the issue of patent quality using different approaches (*e.g.*, Lemley and Sampat, 2012; Frakes and Wasserman, forthcoming). However, they were not designed to quantify the extent of low quality in patent systems.

3. Empirical strategy

Our research seeks to implement the second-pair-of-eye approach with a much larger set of inventions and with more pairs of eyes. Our context allows each patent office to have its own *de facto* standard, and every decision-maker to make mistakes. We do so by analyzing the grant outcome of 'twin' patent applications submitted to multiple jurisdictions. Twin applications are applications covering the same technical content in different jurisdictions (Palangkaraya, Jensen and Webster, 2011; Webster, Jensen and Palangkaraya, 2014; Sampat and Shadlen, 2015).[2] We estimate an index of the probability that each patent application is granted under the differing circumstances of the different patent offices, and then use the resulting estimates to predict the overall ease of obtaining a patent (the threshold) and the proportion of weak patents (inconsistent decisions). The sample for the analysis is the population of 408,133 inventions described in patent applications filed between 2001–2005 in at least two of the EPO (European Patent Office), the USPTO, the JPO (Japanese Patent

[2] Because applicants must submit twin applications to foreign jurisdictions shortly after the submission of the priority filing (up to 12 or 31 months after), the decision to submit twin applications is not driven by the outcome of examination in the office of priority. There is thus no selection on actual grant outcome.

Office), the KIPO (Korean Intellectual Property Office) and the SIPO (State Intellectual Property Office of China). We use this time period in order to ensure that the applicant has had a chance to pursue protection in as many countries as she chooses, and to allow sufficient time to reach a grant decision. These five offices, known collectively as the 'IP5 Offices', attract about 80 percent of worldwide patenting activity.[3]

We employ a reduced-form model of the patent examination decision to separate any systematic factors related to the particular office from the examiner decision about the specific application. Our model of the actual examination decision assumes that each invention has a unique but unobservable inventive merit (c_i), which is therefore shared by all of the applications to different offices. The probability of granting patent application i, by an examiner in office j is a function of this inventive merit c_i (invention fixed effect); the office-specific *de facto* standard required for a grant (τ_j); a set of covariates ($\boldsymbol{x_{ij}}$) capturing observed heterogeneity at the patent-patent office level (*e.g.*, differences in the number of claims, filing route); and examiner-specific factors that are not systematic to the office (ε_{ij}). These elements combine to give an index, y^*_{ij}, which maps into the probability of a grant for each application in each office.

We do not observe this index but rather the binary grant decision, y_{ij}, which takes the value 1 if invention i is granted a patent at office j and 0 otherwise. We estimate y^*_{ij} using a latent variable approach:

$$y^*_{ij} = -\,\tau_j + c_i + \boldsymbol{x_{ij}\beta_j} + \varepsilon_{ij}, \qquad y_{ij} = 1\left[y^*_{ij} > \ 0\right] \tag{1}$$

where a patent for invention i is granted at office j if the latent score is greater than 0. From (1) is can be easily seen that $(-\,\tau_j + c_i)$ is the extent to which the content of the application surpasses the office standard and that $\boldsymbol{x_{ij}\beta_j}$ represents the influence of other systematic features of the office's examination rules. We start by assuming for simplicity that the individual elements of parameter vector $\boldsymbol{\beta_j}$ are constant across j's. In concrete terms, this means that the effect of, *e.g.*, the number of claims on the latent score is common across offices. We will relax that assumption at a later stage.

The stochastic error term ε_{ij} is the aggregation of factors that makes the decision on the criteria for patentability uncertain (*i.e.*, subjective). It captures all of the reasons why, after allowing for the systematic tendencies captured by the regressors, different examiners might reach different decisions on the same invention. That is, if the same application were examined in the same office, under the same office procedures but by a different examiner, any difference in the decision would be explained by ε_{ij}. This term captures, *e.g.*, the subjectivity of interpretation of the patent law or the 'mood' of the examiner. Conceptually,

[3] There were 1,821,150 patent applications filed worldwide in 2010 (priority plus second filings). Of these, 1,452,925 (79.8%) were filed in the IP5 offices (PATSTAT Autumn 2014 version).

if invalidity is only a minor issue, then most of the differences in outcomes at different offices would be due to systematic office effects; in our model this would correspond to the variance of ε_{ij} being small. Conversely, a large variance, causing outcomes across offices to differ even after controlling for invention and office attributes, would be evidence that one or more offices are granting weak patents. An implicit identifying assumption is that $E_j(\varepsilon_{ij}) = 0$, *i.e.*, examiners at office *j* take correct decisions on average. (Any systematic deviation from the 'correct' outcome is captured by the office-specific component.) Likewise $E_i(\varepsilon_{ij}) = 0$, *i.e.*, every invention is treated fairly on average across offices.

We then use the model parameters to tease out the sources of discrepancy in the grant decisions across offices. We call $\widehat{y_{ij}}$ the 'correct' (*i.e.*, predicted) grant outcome and y_{ij} the observed grant outcome. As explained further below, we will estimate equation (1) by means of a linear probability model. That is:

$$\widehat{y_{ij}} = -\tau_j + c_i + \boldsymbol{x_{ij}}\widehat{\boldsymbol{\beta_j}} \qquad (2)$$

The predicted grant outcome is thus based on the linear prediction of the latent quality score (including the invention fixed effect). Since the linear probability model minimizes the mean squared errors, it produces correct inferences on average. This implies that the number of type I errors (mistakenly refused applications) is equal to the number of type II errors (mistakenly granted applications).[4] Hence, patent applications with a predicted latent quality score above the g_j's percentile must be granted, and refused otherwise—where g_j corresponds to the average observed grant rate for applications at office *j*.

We then decompose differences in examination decisions across offices for the same invention (being patent applications that are granted at one office but where the equivalent is refused by at least one other office). As mentioned, this discrepancy has three components: a systematic office effect, capturing *de facto* policies and practices; focal office 'mistake'; and other office 'mistake' (counterpart of a focal office mistake). Recall that patents mistakenly granted at the focal office are what we call weak patents.[5] In practice, we compute the components in the following way:

(a) the grant is 'incorrect' given the focal office's standard for a grant: $y_{ij} = 1$ but $\widehat{y_{ij}} = 0$ ('Focal office mistake', regardless of what the other offices' decisions should have been);

[4] This working assumption may be too strong in light of the argument that it may be rational for the patent office to let bad patents slip through the system (Lemley, 2001). In our model, the fact that examiners may be *systematically* too lenient will be absorbed by the office effect. It is possible to relax this assumption but at the cost of greater computational complexity. Relaxing this assumption would lead to slightly higher rates of weak patents for some offices (not reported).

[5] Some scholars have called 'weak patents' patents that would not stand up in court (Farrell and Shapiro, 2008). In this paper, we call 'weak patents' patents that are at risk of being rejected, should they be re-examiner by the same office. This distinction matters if the courts apply a different standard than the patent office.

(b) the grant is 'correct' given the focal office's standard ($y_{ij} = \widehat{y_{ij}} = 1$) but

(i) the other offices(s) were correct in deciding a refusal: $y_{ik} = \widehat{y_{ik}} = 0$ ('Office effect');

(ii) the other office(s) made a mistake given that their 'correct' decision should be to grant the application: $y_{ik} = 0$ but $\widehat{y_{ik}} = 1$ ('Other office mistake');

We illustrate these various cases in Section 5.

4. Data and variables

4.1 A dataset of one-to-one equivalents across offices

The construction of the dataset is a major undertaking. We combine data from seven offline and online sources. The main data source is the EPO-OECD PATSTAT database (October 2014 release) for the backbone of the dataset. We start from the universe of priority patent applications filed anywhere in the world over the period 2001 to 2005 (de Rassenfosse *et al.*, 2013) and track their one-to-one equivalents in any of the five offices.[6] A priority filing is the first patent application describing an invention. Application P_B in country B is a one-to-one equivalent of application P_A in country A if P_B claims P_A as sole priority (*i.e.*, no merged patent applications) and P_A is only claimed by P_B in office B (*i.e.*, no split patent applications). In this sense, P_A and P_B cover the same technical content and are 'twin' applications. We also extract from PATSTAT information on applicants' country of residence, patents technological fields as identified with the International Patent Classification (IPC) codes, and filing route (either the 'Paris Convention' route or the PCT route).[7]

Data on the application legal status (granted/refused/withdrawn) come from: the EPO's INPADOC PRS table for PATSTAT for European and Chinese applications; from JPO's public access on-line Industrial Property Digital Library Database (IPDLD) for Japanese applications; from KIPO public access on-line IPR Information Service (KIPRIS) for Korean applications; and from the USPTO's Public Pair on-line database for US applications.

Data on the number of claims of published patent applications come from: PATSTAT for European applications; SIPO's on-line patent search platform for Chinese applications; IPDLD for Japanese applications; KIPRIS for Korean applications; and lens.org for US applications. We developed specific web-crawlers to collect online information.

4.2 Variables

---

[6] Thus, our sample may include a priority patent application filed, say, at the Brazilian patent office and with an equivalent at the EPO and the USPTO.

[7] The 'Paris Convention' route is the traditional filing route for patent applications (sometimes call the national route). The term PCT stands for 'Patent Cooperation Treaty.' It is an international treaty that facilitates international patenting.

Our main dependent variable, y_{ij}, is the binary outcome that takes the value of 1 if patent application i was granted by an examiner in patent office *j* and 0 if refused. Our measure of refusal includes applications that were examined and refused by the patent office plus all quasi-refusals. Quasi-refusals include patent applications that were withdrawn at the EPO following a negative search report containing X or Y citations, which challenge the inventive step of an application. Indeed, many applications at the EPO are withdrawn after a (negative) office communication, which Lazaridis and van Pottelsberghe (2007) interpret as quasi-refused applications.

There are three fundamental sources of heterogeneity with respect to the grant outcome in the data: systematic office differences (j), systematic invention differences (i), and application-patent office differences (ij). The first two sources are accounted for by the use of office and invention fixed effects, respectively. Concerning the third source, we control for five variables, $\boldsymbol{x}_{ij}$, that are likely to induce heterogeneity in the grant decision across offices for the same invention. One must really think of variables $\boldsymbol{x}_{ij}$ as control variables that have only a marginal effect on the grant probability. On average the examiners from the different offices make a true assessment of the inventive merit in the application, which is measured with the invention fixed effect. Thus, the four variables influence the examiner's decision over and above the objective quality of the invention.

The first of these controls is a dummy variable, $local\ applicant_{ij}$, which equals 1 if there is at least one applicant with an address in the same jurisdiction as the examining patent office, and 0 otherwise. There is empirical evidence that patent offices give differential treatment to applications based on the country of residence of applicants, with domestic applicants having a higher probability of grant (Webster, Palangkaraya and Jensen, 2014). This home bias may reflect prejudice, but it may also reflect the fact that domestic applicants have stronger incentives to push the patent application in their home market or that they may be more familiar with their home patent system.

The second is the dummy variable $priority\ filing_{ij}$, which takes the value 1 if application *i* is a priority filing in office *j* and 0 otherwise. By the construction of our data (using one-to-one equivalents), there can be only one priority filing per family. Firms usually file a priority filing in the office they know best, which may affect the likelihood that they receive a grant in that office. The country of the priority office may also be the most important market, where incentives to push for a grant are stronger.

The third is the dummy variable PCT_{ij}, which indicates whether the patent application was filed through the Patent-Cooperation Treaty route.[8] There are non-trivial administrative implications of using the PCT route that may affect the consistency of examination outcome

[8] Note that some equivalents are filed partly though the PCT route and partly though the Paris route, leading to within-twin variation.

(*e.g.*, search report shared between all the offices, extension of priority right from 12 to 31 months).

Next, we control for the number of claims ($claims_{ij}$), which is the number of claims articulated in the patent application at the time of lodgment. Although twin applications in our sample cover the same technical content, there might be slight differences in the construction of the applications across offices. The number of claims is a proxy for these differences.

Finally, we include information on the timing of the decision. The variable *Decision #k* takes value 1 for the *k*-th patent application in the family to receive a decision. We expect the probability of grant to decrease with *k*, for two reasons. First, the order of decision could reflect the amount of prior art available to assess the patentability of the invention. In that sense, offices that give a decision later have potentially more prior art available (identified by other offices) to refuse a patent. Second, it could also reflect offices' own judgment about the patent, knowing that it takes longer to refuse a patent application than to accept one. Note that we cannot control for the *nature* of the decision because such variable is correlated with the unobserved invention fixed effect.

Table 1 presents a summary of the characteristics of the patent applications at each office for two samples. The balanced sample (Panel A) is composed of 10,822 inventions for which a patent application has been filed at all five offices (there are thus 54,110 patent applications). The full sample (Panel B) is composed of 408,133 inventions with a patent application in at least two offices, covering in total more than a million applications. Overall, on the full sample, the JPO, at 72.2 percent, recorded the lowest grant rate and the SIPO, at 96.3 percent, the highest. More than half of applications at the JPO had at least one local applicant compared with only 3.1 percent at SIPO.[9] SIPO had also the smallest rate of priority filings and JPO the highest. (Indeed, except for the EPO, there is a strong correlation between the office of priority filing and whether the applicant is local to that office.) Use of the PCT was highest for the EPO but lowest for KIPO. Finally, the average number of claims at the time of application varies between 10.3 at the JPO and 17.8 at the USPTO.

[9] The low proportion of local applicants at the SIPO reflects the fact that very few Chinese firms apply for patent protection in foreign jurisdictions, which is a pre-condition for being in the sample.

Table 1. Descriptive statistics

	N	Grant (%)	local applicant (%)	priority filing (%)	PCT (%)	claims
Panel A. Balanced sample						
EPO	10,822	84.9	27.7	6.3	44.2	14.7
USPTO	10,822	91.5	17.5	18.6	33.0	17.2
KIPO	10,822	88.3	14.7	14.6	4.5	14.9
JPO	10,822	82.6	36.5	36.7	37.7	11.1
SIPO	10,822	97.9	0.6	0.6	21.7	15.2
Panel B. Full sample						
EPO	163,012	76.8	44.2	9.8	45.3	15.6
USPTO	325,068	91.4	20.0	22.3	22.8	17.8
KIPO	127,314	84.4	41.5	41.0	2.3	14.9
JPO	278,760	72.2	56.3	56.4	26.5	10.3
SIPO	170,777	96.3	3.1	3.3	19.7	15.3

Notes: Data relate to patent applications filed between 2000 and 2005. See main text for data sources.

Table 2 provides an overview of the number of equivalents (*i.e.*, twins) between offices. There are 125,704 direct equivalents between the USPTO and the EPO. The lowest number of equivalents is reached between the EPO and the KIPO (32,082 patent applications) and the highest number is reached between the USPTO and the JPO (212,673 applications). As far as the SIPO is concerned, it is most integrated with the USPTO, closely followed by the JPO.

Table 2. Cross-country number of equivalents

	EPO	USPTO	KIPO	JPO	SIPO
EPO	-				
USPTO	125,704	-			
KIPO	32,082	87,228	-		
JPO	91,878	212,673	79,757	-	
SIPO	59,597	119,841	64,925	113,561	-

Notes: Data relate to the full sample.

5. Estimations and results

5.1 Inconsistency rates

We start by examining the data by looking at the 'raw' inconsistency rates, *i.e.*, without correcting for office-specific differences and without neutralizing the influence of examiners' subjective assessments. Results presented in Table 3 show that 28.0 percent of the patents that were granted by the EPO in the balanced sample were refused in at least one other office (21.3% in the full sample). Rates for the balanced sample are logically higher than for the full sample because the probability to observe at least one rejection increases with the number of equivalents that are observed. What matters is that the pattern is similar across both samples: the JPO has always the lowest rates, and the SIPO the highest.

Table 3. Inconsistency rates

	Balanced sample		*Full Sample*	
Office	Number of granted patents	Proportion refused elsewhere	Number of granted patents	Proportion refused elsewhere
EPO	10,822	28.0	125,195	21.3
USPTO	10,822	33.2	297,072	25.2
KIPO	10,822	30.7	107,501	25.7
JPO	10,822	25.9	201,335	13.9
SIPO	10,822	37.5	164,527	26.9

Notes: Data relate to the full sample.

However, as discussed, some of the rejections observed certainly are well founded. The proportion of patents refused elsewhere reflects a combination of systematic differences in policies and practices, mistakes by the focal office and/or mistakes by at least one other office. Next section teases out these sources of heterogeneity.

5.2 Econometric decomposition of the inconsistency rates

There are two conceptually distinct ways to estimate equation (1) econometrically. The first considers that we observe *different outcomes of the same unit* i. The patent examination process is subject to office-specific rules, incentives and biases, and these unobserved factors may or may not be correlated across offices. For example, inventions based on new technologies may be harder to assess against the examination manuals and, therefore, it may be more appropriate to assume $cov(\varepsilon_{ij}, \varepsilon_{ik}) > 0 \; if \; j \neq k$, that is, the omitted explanatory factors for each invention are correlated across offices. Such an approach treats equation (1) as a system of *J* linear equations that one can estimate with a seemingly unrelated regressions (SUR) model. The SUR model has the advantage of taking into account the correlation of errors across offices in the estimation process to improve the efficiency of the estimates. However, implementing fixed effects in a SUR model is not straightforward when the number of individual effects is large. One can control for fixed effects by demeaning the data but at the cost of dropping one equation due to the additivity constraint introduced (leading to a singular variance matrix problem). In addition, the SUR model requires a balanced dataset, which considerably reduces the size of the sample we can use.

The second way considers that we observe the *same outcome in different contexts* j, leading to a fixed-effect (FE) panel data model. The fixed-effect estimator handles unbalanced panels and produces estimates for all offices, which are two desirable features over SUR. However, it does not account explicitly for the fact that the decision errors may be correlated across offices. The extent to which this limitation matters for the present study is an empirical question. As we show below, the SUR and FE models produce quantitatively similar results on the balanced sample—our preferred specification is thus the FE model.

Finally, note that we rely on a linear probability model, which implies that some predicted probabilities might lie outside the unit interval. This issue is of little concern because we are interested ultimately in *ranking* patents by their probability of being granted (and not in the predicted probability score of the grant rate *per se*). In addition, most of the covariates are discrete such that the linear assumption is acceptable. However, we correct standard errors by using heteroskedastic-robust standard errors when appropriate.[10]

We first present results of the econometric model, and then discuss the sources of apparent inconsistency. Table 4 presents the coefficients of equation (1) estimated with different regression models and samples. The column labeled M1 presents an estimate of the SUR model performed on the balanced sample of inventions, having equivalent patent applications at all five offices. As discussed, we need to exclude one office for the model to run, and we arbitrarily exclude the EPO. Column M2 presents results of the fixed-effect estimator for the balanced sample and column M3 for the full sample of inventions with equivalent in at least two jurisdictions. Coefficients in models M1–M3 are constrained to be equal across offices ($\boldsymbol{\beta}$). In model M4, the coefficients for each covariate are office-specific ($\boldsymbol{\beta}_j$), but we only report coefficients for the base group (EPO) for conciseness. Finally, model M5 extends model M4 by controlling for the timing of the decision by offices. The reference group is the office that published the grant (or rejection) decision first.

[10] An alternative estimator is the conditional (*i.e.*, fixed effect) logit estimator. However, it does not exploit information from patent families that are granted or refused at all offices, which is not desirable. Should we use the conditional logit estimator, we would not be able to predict a value for the all negative/all positive outcomes. That is, the predictions we would obtain would be conditional on observing an inconsistency.

Table 4. Determinants of grant outcome

	M1	M2	M3	M4	M5
Regression model:	SUR[(a)]	FE	FE	FE	FE
Sample:	Balanced	Balanced	Full	Full	Full
Coefficients:	Constrained	Constrained	Constrained	Free[(b)]	Free[(b)]
local applicant (LA)	0.126*	0.142*	0.175*	0.138*	0.100*
	(0.007)	(0.006)	(0.002)	(0.002)	(0.002)
priority filing (PF)	0.003	0.018	0.084*	-0.081*	-0.092*
	(0.013)	(0.017)	(0.003)	(0.006)	(0.006)
LA x PF	-0.084*	-0.121*	-0.166*	-0.069*	-0.053*
	(0.016)	(0.019)	(0.004)	(0.006)	(0.006)
PCT	0.034*	0.030*	0.039*	0.127*	0.115*
	(0.004)	(0.004)	(0.001)	(0.003)	(0.002)
claims (log)	-0.007	-0.008	-0.020*	-0.037*	-0.040*
	(0.004)	(0.005)	(0.001)	(0.002)	(0.002)
Timing of decision (ref=1, earliest)					
Decision #2					-0.097*
					(0.001)
Decision #3					-0.148*
					(0.001)
Decision #4					-0.182*
					(0.002)
Decision #5 (latest)					-0.237*
					(0.004)
Office effects (ref=EPO)					
USPTO	0.028*	0.097*	0.176*	0.264*	0.164*
	(0.003)	(0.005)	(0.001)	(0.005)	(0.005)
KIPO	0.007	0.075*	0.123*	0.036*	-0.009
	(0.003)	(0.005)	(0.002)	(0.006)	(0.006)
JPO	-0.074*	-0.004	-0.047*	-0.076*	-0.070*
	(0.003)	(0.006)	(0.002)	(0.005)	(0.005)
SIPO	0.104*	0.172*	0.239*	0.195*	0.165*
	(0.002)	(0.004)	(0.002)	(0.005)	(0.005)
Constant	-	0.821*	0.749*	0.766*	0.890*
		(0.013)	(0.003)	(0.005)	(0.005)
Number of observations	43,288	54,110	1,064,513	1,064,513	1,064,513
Number of inventions	10,822	10,822	408,133	408,133	408,133
R-squared (within)	-	0.053	0.103	0.119	0.153

Notes: * $p < 0.001$; heteroskedastic-robust standard errors in models M2–M5; [(a)] iterated seemingly unrelated regression with demeaned data; [(b)] office-specific coefficients, but only coefficients for the reference group (EPO) reported.

A first observation is that coefficients have similar magnitude and statistical significance between the SUR model (M1) and the FE model (M2), which leads us to adopt the FE model for its flexibility. A second observation is that extending the analysis to the full sample (from model M2 to model M3) produces coefficients that have similar signs but that have stronger statistical significance (expectedly). Notice the strict probability threshold of 1 per thousand for declaring statistical significance of estimated parameters in order to account for the large number of observations. Regarding specific covariates, the results suggest a strong *local applicant* effect, similar to that documented in Webster, Jensen and Palangkaraya (2014). In model M3, the *local applicant* effect is double the magnitude of the *priority filing* effect, and the *local applicant* effect is biggest for non-priority filings. Note that the *priority filing* effect is negative at the EPO (reported in columns M4 and M5) but positive at the other offices (not reported). Patent applications filed through the *PCT* route have a grant rate that is about 3–4 percentage points higher than non-PCT applications (models M1–M3). The effect of the number of claims is always negative, but statistically significant only with the full sample (models M3–M5). Finally, the timing of the decision in model M5 has a strong effect on the probability of grant, with later decisions being systematically less favorable.

Next, we use the estimated parameters of model M5, the most complete model, to tease out the sources of apparent inconsistency in a variance decomposition exercise as explained in Section 3. Let us illustrate the method using the EPO as the focal office. According to Table 3, 21.3 percent of applications granted at the EPO (=26,624) have been refused in at least one other office. We present each of the three cases in turn.

First, a total of 4.0 percent of applications were unduly granted by the EPO. Regardless of whether other offices made a mistake in the applications (in terms of unduly *refusing* an application), these cases correspond to a 'Focal office mistake' and represent what we call weak patents. Second, there are 8.5 percent of applications that were legitimately granted by the EPO and legitimately refused by at least another office ('Office effect'). Third, there are 8.8 percent of applications that were legitimately granted by the EPO (according to the office's own policies and practices) but unduly refused by at least another office ('Other office mistake'). These cases correspond to weak patents being issued by other offices.

Table 5. Correct and incorrect grant at the EPO

	EPO decision = grant		
Other office(s) decision = refusal	*Incorrect grant (% total granted)*	*Correct grant (% total granted)*	
Correct refusal	3,472 (2.8)	10,633 (8.5)	
Incorrect refusal	1,504 (1.2)	11,015 (8.8)	
	4,976 (4.0)	21,648 (17.3)	26,624 (21.3)

Using this method, we can decompose the number of inconsistent grants from Table 3 into the various components for all offices. Doing so leads to the figures presented in Table 6.[11]

Table 6. Decomposition of inconsistency rates, model M5

	Raw rate (Table 3)	Reason(s) for inconsistency: Office effect	Focal office mistake (weak)	Other office mistake
EPO	21.3	8.5	4.0	8.8
USPTO	25.2	15.4	4.0	5.9
KIPO	25.7	10.6	4.8	10.3
JPO	14.0	2.1	5.7	6.2
SIPO	26.9	15.3	1.6	10.0

Notes: The first column corresponds to the last column of Table 3. See main text for details.

Overall, differences in policies and practices across offices account for up to about 15 percent apparent inconsistency at the USPTO and the SIPO and 2.1 percent at the JPO. In other words, the JPO has the highest *de facto* standard and the USPTO and the SIPO the lowest. Mistakes at the focal office (*i.e.*, rate of weak patents) account for as little as 1.6 percent at the SIPO and as much as 5.7 percent at the JPO.

The pattern of ease of patent grant is as would be expected. Japan, the country with the highest standard according to the parameter estimates in Table 4, has a very low rate of granting patents that would be refused by other countries; China has the highest.[12] Of course, we cannot say what is the 'right' standard, so these numbers cannot be strictly interpreted in terms of patent quality. But they do give some quantitative perspective on the possible significance of low standards.

It is tempting to compare the rate of weak patents between offices and conclude that the Chinese patent office is the most 'accurate' office, since it has the lowest figures by this

[11] The results presented in Table 6 consider that the local inventor effect induces legitimate office differences in the grant outcome. Assuming that the local inventor effect is a mistake increases the focal office mistake by a maximum of 0.2 percentage points.
[12] In theory, the strictest office should have a value of 0 in the column *Office effect*. The actual number differs from 0 due to the influence of patent-patent office factors ($\boldsymbol{x}_{ij}$).

measure. However, bear in mind that these figures correspond to *absolute* rates of weak patents, and that one must take into account the fact that offices have varying grant standard. In the limit, if an office has an extremely low standard such that all applications should be granted, it can never make a mistake in the form of granting a patent that it should not have. Conversely, offices with very high standard have more room for making mistakes. One can normalize the figures by estimating how much the office decision deviates from a random decision-making using the observed grant rate. For example, knowing that the observed grant rate at the EPO for the full sample is 76.8 percent, a random grant decision would produce 17.8 percent of Type I and 17.8 percent of Type II errors (76.8×(1-0.768) and 23.2×0.768). Relative to the total proportion of granted patents (0.768), the invalidity rate of random decisions would be simply 1-0.768 = 0.232. Since the estimates imply that the EPO made 'only' 4.0 percent of Type II errors, its relative accuracy is 0.232/0.04 = 5.8. The interpretation is straightforward: should the EPO take random grant decisions, it would grant 5.8 times as many weak patents as it currently does. That is, the rate of weak patents is about 17 percent of the random error rate for the EPO. The relative accuracy rates at the other offices are 2.15 (USPTO), 3.25 (KIPO), 4.8 (JPO) and 2.3 (SIPO), which implies that the EPO and JPO are the most accurate offices and the USPTO and SIPO the least accurate.

6. Discussion and robustness tests

6.1 Accounting for differences in patentable subject matters

Although the empirical analysis controls for five covariates that are likely to induce heterogeneity in the application-patent office pair, one potential source that is not accounted for is the difference in patentable subject matter across jurisdictions. Such differences would lead to a legitimate grant at one office and a legitimate refusal at another office, but according to our method, would be interpreted as an error in one office.

We know from discussions with patent attorneys that the definition of patentable subject matter in mechanical engineering is very similar across jurisdictions. Hence, we can use the field of mechanical engineering as a benchmark for errors that are not affected by difference in patentable subject matter definition.

Table 7 assigns each family to one or more major technology OST technology groups based on any one of the IPC subclasses given at any office.[13] In addition, we use the 'Biotechnology' and 'Software' classifications from the OECD (2003) and Graham and Mowery (2004) respectively. Table 7 breaks down the inconsistency rates by technology field. The estimates are based on model M5, that is, the fixed-effect estimator with office-specific coefficients run on the full sample and controlling for the timing of office decision.

[13] Office of Science and Technology, UK classifications.

Table 7. Proportion of weak patents by technology fields, model M5.

	EPO	USPTO	KIPO	JPO	SIPO
Electrical	5.3	4.2	4.2	5.8	1.8
Instruments	4.4	4.0	5.1	5.6	1.7
Chemicals & pharmaceuticals	3.1	6.3	6.0	7.2	1.6
Process engineering	3.5	5.0	5.4	6.1	1.4
Mechanical engineering	**3.0**	**3.3**	**5.4**	**4.9**	**1.2**
Biotechnology†	3.7	7.3	6.3	7.9	2.6
Software††	6.4	6.3	4.7	7.7	2.2

Notes: An application can be allocated to more than one major technology groups based on multiple IPC subclasses assigned in any office. Major OST group excluding Biotechnology and Software. †Based on OECD (2003). ††Based on Graham and Mowery (2004).

One can read the results in Table 7 in two ways. First, if one believes that differences in patentable subject matter across offices affect the estimates presented in Table 6, then one should only focus on the estimates for the field of mechanical engineering. The rates of weak patents are slightly lower than those presented in Table 6 but the ranking across countries is globally consistent. Thus, concerns that differences in the patentable subject matters may drive results seem misplaced. Second, if one believes that there are no major differences in patentable subject matters across offices for applications in our sample, then the estimates can be taken as reflecting differences in rates of weak patents across fields.[14] Qualitatively, offices have a relatively high rate of weak patents in software and in biotechnology. This pattern is loosely consistent with the notion that subjective judgment about patentability is harder in these newer fields.

We do not report the values of the office effects for conciseness. However, we briefly comment the result obtained for biotechnology patents for the EPO and the USPTO. Whereas the USPTO effect is estimated at 0.164 relative to the EPO in the full sample (model M5), the effect for the sample of biotechnology patents is considerably higher, at 0.293. This pattern is consistent with the discussion in Hopkins *et al*. (2007), which explains that the EPO has taken a more stringent approach than the USPTO on DNA-relation inventions.

6.2 External validity

Patents applications in our sample are considerably less selected than in litigation studies previously used to study invalidity. Compared to previous studies, the sample does not select on likely (in)validity. Our sample does, however, select on invention economic value, because applicants are more likely to pursue protection in multiple countries for more valuable inventions. Although patent value is not a patentability requirement, we cannot exclude the

[14] We do not expect to find that many focal office mistakes are traced to differences in patentability subject matter as the bulk of our sample is composed of experienced applicants who would not file patent applications in jurisdictions where the subject matter was not patentable.

possibility that economic value may be correlated with invention quality and we therefore investigate the extent of selection in the data.

A first selection that might occur is selection on quality with respect to the filing decision, that is: Are higher quality inventions more likely to be filed abroad (and hence more likely to appear in our sample)? One way of testing for the presence of selection at office *j* involves estimating equation (1) for all offices but *j* and assessing whether the recovered invention fixed effect (*i.e.*, estimated quality) predicts filing at office *j*. Table 8 reports the mean value of the fixed effect thus computed by filing status at each office. In the first row, we obtain the invention fixed effect by estimating equation (1) with ignoring EPO observations. We then compute the mean score of the fixed effect by filing status (filed/not filed) at the EPO. Overall, the results suggest that quality does affect the filing decision, with higher quality patents being more likely to be filed in foreign jurisdictions.

The last column of Table 8 reports the marginal effect at the mean of a one-standard deviation increase in quality on the filing decision. For instance, a one-standard deviation increase in invention quality leads to a 3.7 percent increase in the probability that a patent application will be filed at the EPO. Selection is strongest at the USPTO and weakest at the EPO. Thus, it appears that our sample is biased to a small but not trivial extent towards inventions with higher than average quality.

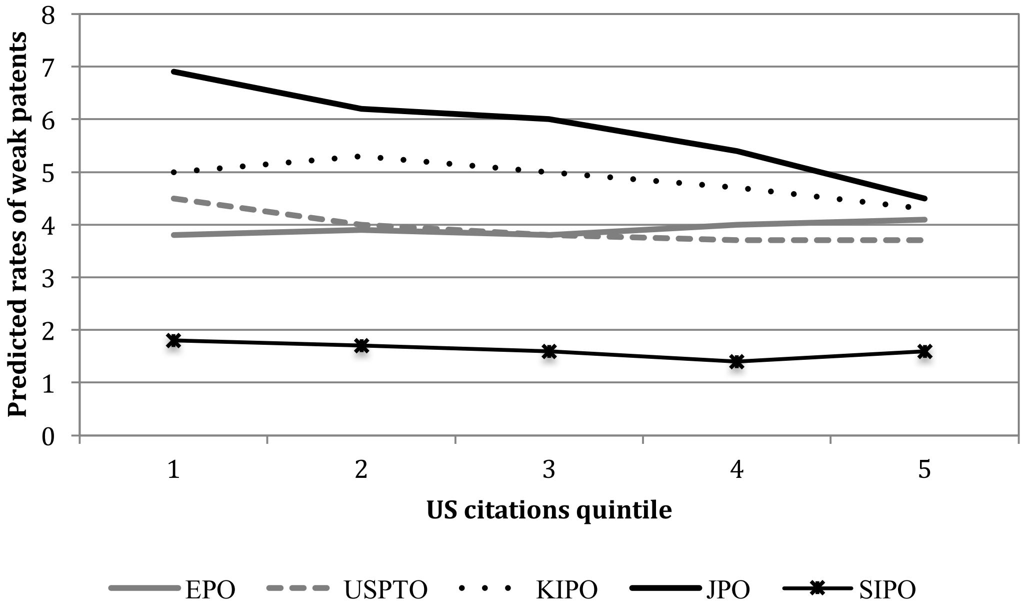
Table 8. Invention quality by filing status

	Not filed	Filed	Δ	Marginal effect
EPO	-0.019	0.016	-0.035*	0.037
USPTO	-0.121	0.045	-0.165*	0.115
KIPO	-0.025	0.030	-0.055*	0.052
JPO	-0.042	0.021	-0.062*	0.064
SIPO	-0.042	0.053	-0.096*	0.087

Notes: Columns 'Not filed' and 'Filed' report the mean score of the invention fixed effect and 'Δ' is the difference. *: p<0.001.

The fact that inventions in our sample are somewhat selected on their quality does not tell us anything directly about possible bias in our estimates. We assess the effect of quality on the rate of weak patents by relying on a commonly used quality indicator, namely the number of forward citations. As recently reviewed by Jaffe and de Rassenfosse (2016) there is a long tradition in the literature of using forward citations to proxy the technological merit of the invention (Albert et al., 1991; Narin, 1995; Trajtenberg, Henderson and Jaffe, 1997). Figure 1 presents the relative rates of weak patents by quintiles of citations received at the USPTO. We count citations received by USPTO patents from USPTO patents up to seven years after first publication using the PATSTAT database (de Rassenfosse, Dernis and Boedt, 2014:402). Overall, the proportion of weak patents seems to decrease with the number of citations received, especially at the JPO, where rates of weak patents go down from 7 percent to less than 5 percent.

Figure 1. Proportion of weak patents by citations received



Notes: 0 citation for the first quintile; Q2: 1 citation; Q3: 2 or 3 citations; Q4: 4 or 5 citations; Q5: 6 citations or more.

Summarizing the insights from both tests we come to the following conclusions. Selection into filing at the EPO is small and the effect of quality on the rate of weak patents is stable across quintiles of the quality metric. Therefore, the population-wide rate of weak patents is likely to be around 4 percent. A similar reasoning holds for the KIPO, with a population-wide rate of weak patents of about 5 percent. There is strong selection into filing at the SIPO but the rate of weak patents is fairly stable across quintiles of the quality metric such that the population-wide figure is probably close to 2 percent anyway. In light of the strong selection into the filing decision at the USPTO, the population wide rate of weak patents is probably closer to 5 percent than 4 percent. At the JPO, population-wide rate is probably closer to 7 percent than 5 percent for similar reasons.

6.3 Sensitivity to applicant experience

Legal scholars argue that patent prosecution is fundamentally a negotiation between the applicant and the examiner (*e.g.*, Brunner, 2014). This observation suggests that applicant's experience may interfere with our estimates. On the one hand, more experienced applicants are presumably better equipped to push their patents through the examination process. On the other hand, more experience applicants may spend less energy in each application, leading to potentially more heterogeneity in grant decisions. We investigate whether the rate of weak patents varies with the level of experience of applicants.

Figure 2. Proportion of weak patents by applicant experience

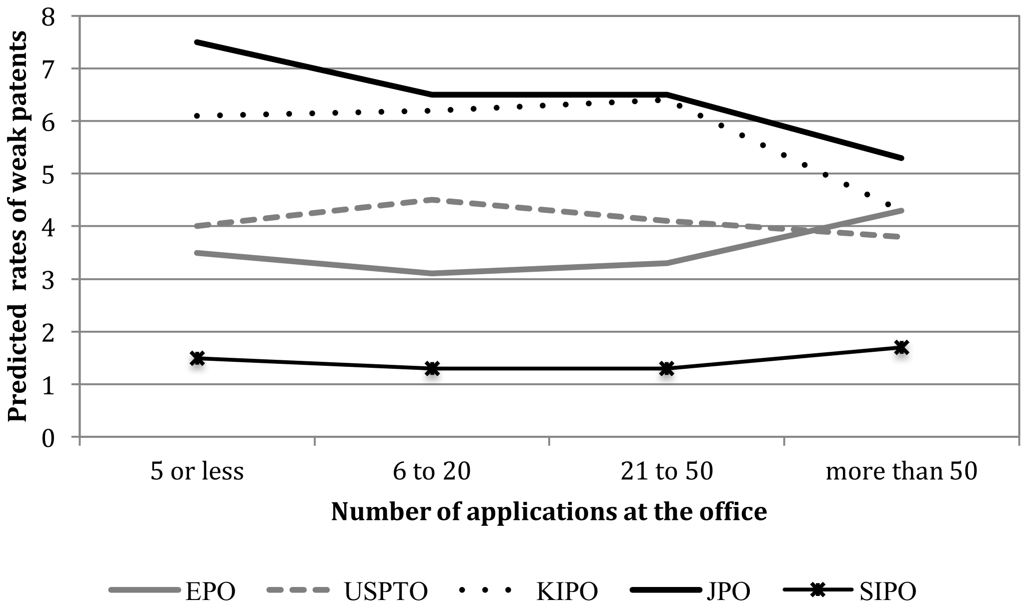


Figure 2 depicts the rates of weak patents by applicant experience (measured in terms of the number of applications submitted to the focal office over the whole study period). Overall, no clear pattern emerges.

6.4 Additional considerations

One source of unobserved heterogeneity relates to the scope of protection for the same invention across offices. Two offices may grant a patent yet one office may be more stringent than the other by limiting the scope of the claimed invention. It is reasonable to argue that leniency in the scope of protection will also translate into higher issuance rate, such that the binary outcome that we observe should lead to correct inference about the office effect. However, our method may be underestimating the rate of weak patents, and even more so for the most lenient offices. Given that the patent applications in our sample are written in four or five different languages, it is extremely difficult to compare the scope of protection across offices. However, we can restrict the sample to a set of highly homogeneous applications to get a sense of the severity of the issue. We have estimated model M5 on the subsample of 322,583 applications with the same number of claims at filing across jurisdictions. Doing so gives qualitatively similar results (not reported).

Finally, there is some question about whether the PATSTAT database correctly records all Japanese language PCT applications to the JPO that were refused. We find no evidence that these applications are missing from the central PATSTAT file. However, to accommodate the possibility that these applications are erroneously tagged as pending, we recoded as ‘refused’ all Japanese applicants who filed at the JPO through the PCT but have no recorded legal status. This amounted to 36 applications and did not change the results.

7. Conclusion

There is significant concern around the world that patent offices are issuing patents that should not have been granted. Studies based on litigation outcomes suggest that this problem is quantitatively significant, with the overall fraction of dubious patents ranging from a quarter to three-quarter of all patents. Our analysis of patent applications examined by multiple offices around the world suggests that the overall prevalence of low-quality patents is likely to be smaller.

We model the patent grant process in a way in which imperfect decision-makers compare their assessment of the quality of an invention to an internal standard of quality necessary for grant. This method allows us to decompose differences in the decisions of multiple decision-makers into those that are due to a mistake by a decision-maker and those that are due to differences in the policies and practices applied by different decision-makers (office specific differences). The kind of decomposition that we have undertaken requires repeated observations on each invention and each decision-making unit.

Our analysis of about 400,000 inventions considered for patent protection by multiple patent offices suggests that both sources of inconsistent decisions are important. The strength of our analysis is to compare various offices using the same data. To push the point, it allows us to conclude that differences in grant outcomes are primarily driven by policy choices and practices rather than subjectivity of the examination process.

Specifically, we find that the fraction of weak patents—those that should not have been granted given the offices own grant standard—does not exceed single digits for any office. Having noted this, we find that some offices are better at screening patent applications than others. Our data suggest that the examiner decisions at the USPTO and the SIPO are twice as accurate as a random decision, whereas the EPO and the JPO are five times more accurate.

While the sample used for the analysis is large, it is not randomly drawn. Patents examined in multiple international jurisdictions are likely to be of higher economic value than the average patent. Our analysis of the selection problem suggests, however, that rates of weak patents for the population of all applications to each office are unlikely to be much higher than our estimates for this IP5 sample. Thus, even allowing for selection bias, our results suggest rates much lower than the rates found by litigation studies.

The (much) lower rates of weak patents obtained with our method compared to litigation studies can be explained by four factors. First, litigated patents are highly selected towards those most likely to be found invalid. Second, litigation studies implicitly assume that courts apply the same standard as that of the office whose grant is being reviewed, and do not make mistakes themselves. In practice, it is possible that courts systematically apply a stricter standard for validity than the patent office—and make mistakes themselves. Third, although patent applications in our sample are examined by up to five examiners from very different cultures and language groups, every examiner spends considerably less time than if the patent were re-examined in litigation. Finally, review by a court is fundamentally different from review by another examiner because the court review is an adversarial proceeding. It is possible that there is prior art that no patent examiner will ever find, but which the adverse party is able to bring to the court's attention. Thus overall our results provide a different perspective on patent quality and should be viewed as complementary to those of litigation studies rather than directly comparable.

The magnitude of the difference between the figures presented in this paper and the figures obtained using patent litigation data bear important implications for discussions about patent quality. One difficulty in interpreting the difference is that we do not know how much of it might be due to selection bias in the litigation studies. But if we assume for the sake of argument that invalidity in the view of the courts is truly significantly higher than invalidity in the view of the offices, we can make four general points. First, much of the debate around quality focuses on improving examination. Our results suggest that this effort is somewhere between misguided and only marginally useful. Second, some of the debate has a flavor of the United States, in particular, having a low standard. Our results suggest that while it is true that the U.S. standard is somewhat low, raising it to the level of the highest country would have only a modest impact. Third, more generally, the tone of the debate is frequently that the uncertainty around validity is the patent offices' fault. Our results suggest rather that it is inherent in the examination process that a non-trivial number of invalid patents will be approved. Finally, we bring into sharp focus the question of *why* courts are more likely to invalidate than examiners. To the extent that it is because of the adversarial nature of litigation, the finding brings the question of how to best to organize re-examination processes that are undertaken within offices. But if it is because judges are fundamentally tougher than examiners, the finding raises deeper questions about administrative law, since judges are not supposed to apply different standards.

The findings presented in this paper are interesting in their own right in light of concerns about patent quality, but they also contribute to current policy discussions on patent prosecution highway (PPH) agreements. PPH designates a set of initiatives for providing accelerated prosecution procedures by sharing information between patent offices. Our results show that there is considerable heterogeneity across offices. The PPH agreements intend to increase the harmonization of decision. However, they may also propagate a wrong decision into the whole patent family, further weakening patent rights. Our results further illustrate that some offices are more accurate than others, which may create additional tensions in the context of PPH agreements.

The fraction of patents that might be said to be low quality in the sense that they result from systematically lax policies and practices is larger than the rate of weak patents, ranging from 9 percent for the EPO to approximately 11 percent for Korea and 15 percent for the United States and China. Even these larger numbers seem modest relative to the general policy discussion about the problem of patent quality. This suggests some notion that all of these offices have a grant standard that is too low relative to some normative judgment. Our analysis—based as it is on comparing decisions across offices—sheds no light on that issue.

Finally, our analysis is silent on the optimal level of ambiguity that the patent system should tolerate. On the one hand, weak patents hurt businesses and may slow down the pace of technological progress. On the other hand, ensuring high quality examination is costly, especially in light of the fact that the majority of patents have limited economic potential. Future research should investigate whether delivering more harmonized outcomes for businesses is likely to improve welfare. Our results provide a useful starting point in that regard.

Acknowledgments

Dan L. Burk, David Card, Annamaria Conti, Dietmar Harhoff, Joachim Henkel, Karin Hoisl, Sonia Jaffe, Keld Laursen, Yann Ménière, Arti Rai, Ben Roin and Carl Shapiro provided valuable comments. The authors are also grateful to seminar and conference participants at the NBER Summer Institute, Duke Law School, Congress of the European Economic Association, European Patent Office, European Association for Research in Industrial Economics Conference, European Policy for Intellectual Property Conference, Toulouse School of Economics, ETH Zurich, the third International Meeting in Law & Economics (Paris), the Munich Summer Institute, New Zealand Economic Association Conference, and Asia Pacific Innovation Conference, T'Mir Julius provided excellent research assistance and her contribution is gratefully acknowledged. This study was financed by the Australian Research Council Discovery Grant ARC LP110100266 'The Efficiency of the Global Patent System' with partners IP Australia and the Institute of Patent and Trademark Attorneys.

References

Albert, M. B., Avery, D., Narin, F., & McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research policy*, *20*(3), 251–259.

Allison, J., & Lemley, M. (1998). Empirical evidence on the validity of litigated patents. American *Intellectual Property Law Association Quarterly Journal*, 26(3), 185–275.

Allison, J., & Hunter, S. (2006). On The Feasibility Of Improving Patent Quality One Technology At A Time: The Case Of Business Methods. *Berkeley Technology Law Journal*, 21, 729-794.

Brunner, J. (2014). Patent Prosecution as Dispute Resolution: A Negotiation Between Applicant and Examiner. *Journal of Dispute Resolution*, 3(1), 7–21.

Cremers, K., Ernicke, M. Gaessler, F. Harhoff, D. Helmers, C. McDonagh, L. Schliessler, P. Van Zeebroeck, N. (2013) Patent litigation in Europe, *ZEW Discussion Paper* 13-072.

Choi, J. P., & Gerlach, H. (2016). Patent pools, litigations, and innovation. *RAND Journal of Economics*, forthcoming.

de Rassenfosse, G., Dernis, H., & Boedt, G. (2014). An introduction to the Patstat database with example queries. *Australian Economic Review*, *47*(3), 395-408.

de Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., & van Pottelsberghe, B. (2013). The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy*, *42*(3), 720–737.

de Saint-Georges, M., & van Pottelsberghe, B. (2013). A quality index for patent systems. Research Policy, 42(3), 704–719.

The Economist (2015). A question of utility. August 8 2015, pp 50-52.

Encaoua, D., & Lefouili, Y. (2009). Licensing 'weak' patents. *Journal of Industrial Economics*, 57(3), 492–525.

Farrell, J., & Shapiro, C. (2008). How strong are weak patents? *American Economic Review*, 98(4), 1347–1369.

Frakes, M. & Wasserman, M. (forhtcoming). Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Micro-Level Application Data, *Review of Economics and Statistics*, in press.

Graham, S. & Mowery, D. (2004). Submarines in software? continuations in US software patenting in the 1980s and 1990s. *Economics of Innovation and New Technology*, 13, 443–456.

Hall, B., Jaffe, A., & Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. *NBER Working Paper* 8498.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg (2005). Market value and patent citations. *RAND Journal of economics*: 16-38.

Henkel, J. and Zischka, H. (2014) 'Why most patents are invalid – Extent, reasons, and potential remedies of patent invalidity' mimeo, TUM School of Management, Technische Universität München, 29 September 2014.

Jaffe, A. B. and de Rassenfosse, G. 2016. Patent Citation Data in Social Science Research: Overview and Best Practices. *Journal of the Association for Information Science and Technology*, forthcoming.

Jaffe, Adam B. and Josh Lerner (2004). Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It. Princeton: Princeton University Press.

Hopkins, M. M., Mahdi, S., Patel, P., & Thomas, S. M. (2007). DNA patenting: the end of an era?. *Nature biotechnology*, 25(2), 185–187.

Lazaridis, G., & van Pottelsberghe, B. (2007). The rigour of EPO's patentability criteria: An insight into the 'induced withdrawals.' *World Patent Information*, 29(4), 317–326.

Lemley, M. A. (2001). Rational ignorance at the patent office. *Northwestern University Law Review* 95.4.

Lemley, M. A., & Sampat, B. (2012). Examiner Characteristics and Patent Office Outcomes. *Review of Economics and Statistics*, 94(3), 817–827.

Lemley, M. A., & Shapiro, C. (2005). Probabilistic Patents. *Journal of Economic Perspectives*, 19(2), 75–98.

Marco, A. (2004). The selection effects (and lack thereof) in patent litigation: Evidence from trials. The B.E. Journal of Economic Analysis & Policy, 4(1), 1226.

Merges, R., & Nelson, R. (1990). On the complex economics of patent scope. Columbia Law Review, 90(4), 839–916.

Meurer, M. J. (2009). Patent Examination Priorities. *William & Mary Law Review*, 51(2), 675–709 .

Miller, S. (2013). Where's the innovation? An analysis of the quantity and qualities of anticipated and obvious patents. *Virginia Journal of Law and Technology*, 18(1), 1–58.

Nagaoka, S. and Yamauchi, I. (2015). Information constraint of the patent office and examination quality: Evidence from the effects of initiation lag. Mimeo.

Narin, F. (1995). Patents as indicators for the evaluation of industrial research output. *Scientometrics*, 34(3), 489–496.

OECD (2003) Science, Technology and Industry Scoreboard, Organization for Economic Cooperation and Development, Paris.

Palangkaraya, A., Webster, E. & Jensen, P. (2011). Misclassification between patent offices: Evidence from a matched sample of patent applications. *Review of Economics and Statistics*, 93(3), 1063–1075.

Paradise, J., Andrews, L., & Holbrook, T. (2005). Patents on Human Genes: An analysis of Scope and Claims. *Science*, 307, 1566–1567.

Picard, P., & van Pottelsberghe, B. (2013). Patent office governance and patent examination quality. *Journal of Public Economics*, 104, 14–25.

Sampat, B. & Shadlen, K. (2015). TRIPS implementation and secondary pharmaceutical patenting in Brazil and India*. Studies in Comparative International Development*, 50 (2), 228–257.

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

Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1), 19–50.

Wallach, E. J., & Darrow, J. J. (2016). Federal Circuit Review of USPTO Inter Partes Review Decisions, by the Numbers. *Journal of the Patent & Trademark Office Society*, *98*, 105.

Webster, E., Jensen, P., & Palangkaraya, A. (2014). Patent examination outcomes and the national treatment principle. *RAND Journal of Economics*, 45(2), 449–469.