Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13-15, 2016

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

Gaétan De Rassenfosse
EPFL
CDM
gaetan.derassenfosse@epfl.ch

Adam Jaffe
Motu Economic and Public Policy Research
NA
Adam.Jaffe@motu.org.nz

Elizabeth Webster
Swinburne University of Technology
Centre for Transformative Innovation
emwebster@swin.edu.au

Abstract

Low-quality patents generate business uncertainty and may create unjustified monopoly rights. There presence is of considerable concern to businesses operating in patent-dense markets. We define low-quality patents as patents whose inventive step is too small to deserve patent protection. There are two pathways by which this may be occurring: the patent office may apply systematically a standard that is too lenient (low inventive step threshold); and the patent office may grant patents that are, in fact, below the threshold (so-called 'weak' patent). This paper uses novel data from 'twin' patents that have been examined at the five largest patent offices to derive first-of-their-kind office-specific estimates of the height of the inventive step threshold and the prevalence of weak patents.

Jelcodes:O34,L43

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

This version: February 25th, 2016

Abstract

Low-quality patents generate business uncertainty and may create unjustified monopoly rights. There presence is of considerable concern to businesses operating in patent-dense markets. We define low-quality patents as patents whose inventive step is too small to deserve patent protection. There are two pathways by which this may be occurring: the patent office may apply systematically a standard that is too lenient (low inventive step threshold); and the patent office may grant patents that are, in fact, below the threshold (so-called 'weak' patent). This paper uses novel data from 'twin' patents that have been examined at the five largest patent offices to derive first-of-their-kind office-specific estimates of the height of the inventive step threshold and the prevalence of weak patents.

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 commercialisation 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 patent system relates to the fact that patent offices might *apply systematically 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, scholars have documented flaws in the examination process. There is empirical evidence that the likelihood of grant is affected by the amount of information that is available to the examiner about the relevant prior art (Nagaoka and Yamauchi, 2015), by the experience of examiners (Lemley and Sampat, 2012) and by the time available for examination (Frakes and Wasserman, 2014). More generally,

[1] See http://www.uspto.gov/patent/initiatives/patent-quality-chat

the grant decision rests ultimately on a subjective comparison of the application's inventive step 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. Such uncertainty only comes about if standards are not applied consistently. Scholarly literature refers to inconsistent patents as 'weak' patents and shows 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 in which these two effects are both present (low inventive step threshold and weak patents). We then use novel data on multiple examination outcomes for the same invention in different patent offices to estimate their magnitudes. 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 quality of the patent grant secured elsewhere. In particular, we estimate a statistical model of the grant process that captures parametrically the effects of different grant standards in different countries, the effect of observable application attributes on the grant probability, and the possibility of personal (*i.e.*, examiner) discretion in every decision. The estimated parameters are then used to quantify the height of the inventive step and the prevalence of weak patents.

To foreshadow the results, we find that differences across offices in the threshold appear to be quantitatively more significant than within-country inconsistency of decisions, but such inconsistency is present to varying degrees across countries. The model estimates imply that only 2–6 per cent 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. An additional 2-15 per cent can be thought of as low-quality in the sense that they would not have been granted if the patent standard of some other country had been applied. The inventive step threshold appears to be lowest in China and the United States, and highest in Japan. The rate of weak patents does not exceed single digits for all offices.

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 determinants of invalidity. However, they cannot distinguish 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.

Patent litigation studies report 'invalidity' rates in the range between 30 to 75 per cent. 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 per cent 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 per cent 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 per cent 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.

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 per cent. 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-per cent figure may still be biased, though the direction of bias is unclear.

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 per cent) were problematic. The 'second-pair-of-eyes review' program at the USPTO, which began in the year 2000, aims at assessing examination quality by re-examining a random set of patent applications. 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."

3. Empirical strategy

Our research seeks to implement the second-pair-of-eye approach with a much larger set of inventions and with many pairs of eyes. Our context allows each institution to have its own implicit standard, and every decision-maker to make mistakes. We do so by analysing the grant outcome of 'twin' patents applications submitted to multiple jurisdictions. Twin applications are applications covering the same technical content in different jurisdictions.[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 inventive step threshold and the proportion of mistakenly granted (*i.e.*, weak) patents. 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 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 per cent 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 step (c_i), and this inventive step 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 step c_i (invention fixed effect); the office-specific inventive step threshold 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, whether the application is filed through the PCT); and examiner specific factors that are not systematic to the office. These factors 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:

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

[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 (revealed) grant decision in the office of priority. There is thus no selection on actual grant outcome.

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

where a patent for invention i is granted at office j if the latent score is greater than 0. We start by assuming for simplicity that the individual elements of parameter vector $\boldsymbol{\beta}$ are constant across j's. In concrete terms, this means that the effect of, say, the number of claims on the grant outcome 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}. Conceptually, 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 invalid patents. We use this information to quantify the rate of invalidity. 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 'true' outcome is captured by the office-specific component.) Likewise $E_i(\varepsilon_{ij}) = 0$, *i.e.*, every invention is treated fairly on average.

We then use the model parameters to tease out the sources of 'apparent inconsistency' in the grant decisions across offices. We call $\widehat{y_{ij}}$ the true (predicted) grant outcome and y_{ij} the observed grant outcome. The apparent inconsistency is composed of patent applications that are granted at one office but where the equivalent is refused by at least one other office. We call it the 'raw invalidity' rate. Conceptually, one can decompose this rate into three components: focal office mistake; office inventive step threshold difference; and other office mistake. In practice, we compute them in the following way:

(a) the grant is 'true' given the focal offices inventive threshold but
    (i) the other office(s) made a mistake given that their 'true' decision should be to grant the application ('other office mistake');
    (ii) the other offices(s) were correct in deciding a refusal ('office threshold differences');
(b) the grant is 'false' given the focal offices inventive threshold and
    (i) The other office(s) were correct in deciding a refusal ('focal office mistake');
    (ii) the other office(s) were mistaken in deciding a refusal as their 'true' decision should be a grant ('office threshold differences').

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. It combines 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.[4] Application B is an equivalent of application A if B claims A as sole priority (*i.e.*, no merged patent applications) and A is only claimed by B in B's office (*i.e.*, no split patent applications). In this sense, A and 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 'Patent Cooperation Treaty', PCT, route).

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

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.[5] 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

[4] 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.

[5] In reality, there is a spectrum of possible examination outcomes. In particular, an application may be granted but have some of its claims denied in one or more offices. We have not explored the empirical significance of this possibility. Differences in languages of the patent documents across offices make such an approach challenging to implement in a large scale.

(negative) office communication, which Lazaridis and van Pottelsberghe (2007) take as evidence of quasi-refused applications.

There are three observable 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 four variables, $\boldsymbol{x}_{ij}$, that are likely to induce heterogeneity in the grant decision across applications for the same 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 clear 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, 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. The PCT is an international treaty that facilitates international patenting. There are non-trivial administrative implications of using the PCT route that may affect the consistency of examination outcome (*e.g.*, search report shared between all the offices, extension of priority right from 12 to 31 months).

Finally, 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, the actual scope of the application may differ across offices. The number of claims is a proxy for differences in the scope of protection.

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 per cent, recorded the lowest grant rate and the SIPO, at 96.3 per cent, the highest. More than half of applications at the JPO had at least one local applicant compared with only 3.1 per cent at SIPO.[6] SIPO had also the smallest rate of priority filings and JPO the highest. (Indeed, 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.

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, 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.

[6] The low proportion 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.

5. Estimations and results

5.1 Raw invalidity rates

We start by examining invalidity by looking at the 'raw' invalidity rates, *i.e.*, without correcting for office-specific differences and without neutralising the influence of examiners' subjective assessments. Results presented in Table 3 for the full sample of patent applications show that 21.3 per cent of the patents that were granted by the EPO were refused in at least another office. The corresponding figure is highest at the SIPO, where 26.9 per cent of patents granted were refused at least once elsewhere and lowest at the JPO, with a rate of 13.9 per cent.

Table 3. Raw invalidity rates

Office	Number of granted patents	Proportion refused elsewhere
EPO	125,195	21.3
USPTO	297,072	25.2
KIPO	107,501	25.7
JPO	201,335	13.9
SIPO	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 legitimate inventive step threshold differences, 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 estimates of the invalidity rates

There are two conceptually distinct ways to estimate equation (1). 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 predicted invalidity rates are very similar between both the SUR and FE models—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.[7]

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 labelled 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.

[7] An alternative estimator is the conditional logit estimator. However, we cannot use information from patent families that are granted at all offices, which is not desirable.

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). 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 thousands 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 has a strong effect on the probability of grant, with later decisions being systematically less favourable.[8]

Next, we use the estimated parameters (model M5) to tease out the sources of apparent inconsistency. As explained, we first need to compute the 'true' outcome $\widehat{y_{iJ}}$ for all patent applications in our sample. We then break down the raw invalidity rate as explained in Section 3. Let us illustrate the methodology with an example using the EPO as the focal office. According to Table 3, 21.3 per cent 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, there are 8.8 per cent of applications that were rightly granted by the EPO (we observe a grant at the EPO, and we also predicted a grant at the EPO) and we also predicted that the applications should have been granted at all other offices. Since at least another office refused the patent, it means that another office than the EPO made a mistake.

Second, there are 8.5 per cent of applications that were wrongly granted at the EPO (we observe a grant at the EPO, but predicted a refusal at the EPO) and we also predicted that the applications should have been granted at all other offices. Hence, these cases must be explained by a difference in office threshold between the EPO and the other offices.

Third, there are 4 per cent of applications (2.8 + 1.2) that we observe as granted at the EPO, yet we predict that they should have been refused at least at another office (and in fact, they were), accounting for how lax/strict offices are. These sources of errors reveal an inconsistent decision from EPO examiner.

[8] There are two potential reasons for this negative effect. 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. It could also reflect offices' own judgment about the patent, knowing that it takes longer to refuse a patent application than to accept one.

Table 5. True and false grant at the EPO

Other office(s) decision = refusal	EPO decision = grant		
	False grant (% total granted)	*True grant (% total granted)*	
True refusal	3,472 (2.8)	1,504 (1.2)	
True grant	10,633 (8.5)	11,015 (8.8)	
	14,105 (11.3)	12,519 (10.0)	26,624 (21.3)

Using this method, we can decompose the number of raw invalid grants from Table 3 into the various components for all offices. Doing so leads to the invalidity rates presented in Table 6.

Table 6. Decomposition of raw invalidity rates, model M5

	Raw rate (Table 3)	Sources		
		Difference in office threshold	Focal office mistake	Mistakes at other offices
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 office threshold account for up to about 15 per cent apparent inconsistency at the USPTO and the SIPO and 2.1 per cent at the JPO. In other words, the JPO has the highest threshold and the USPTO and SIPO the lowest. Mistakes at the focal office (*i.e.*, rate of weak patents) account for as little as 1.6 per cent at the SIPO and as much as 5.7 per cent at the JPO.

The pattern of low grant thresholds is as would be expected. Japan, the country with the highest threshold 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.[9] 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 thresholds.

Turning back to the issue of invalidity in the sense of internal inconsistency, 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 invalidity rate by

[9] In theory, the strictest office should have a value of 0 in the column *Difference in office threshold*. The actual number differs from 0 due to the influence of patent-patent office factors ($\boldsymbol{x}_{ij}$).

this measure. However, bear in mind that these figures correspond to *absolute* invalidity rates, and that one must take into account the fact that offices have varying grant thresholds. In the limit, if an office has an extremely high threshold such that all applications should be refused, it can never make a mistake in the form of granting a patent that it should not have. Conversely, offices with very low thresholds have plenty of room for making mistakes because there are a lot of refused applications. One can normalize the invalidity rates 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 per cent, a random grant decision would produce 17.8 per cent of Type I and Type II errors (0.768×(1-0.768)). 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 per cent 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 invalid patents as it currently does. The relative accuracy rates at the other offices are 2.15 (USPTO), 3.25 (KIPO), 4.8 (JPO) and 2.3 (SIPO).

6. Discussion and robustness tests

6.1 Accounting for differences in patentable subject matters

Unobserved heterogeneity in our model takes the form of systematic patent-patent office effects (c_{ij}). Such effects fall into the error term and affect the invalidity rates. Although the empirical analysis controls for four covariates that are likely to induce heterogeneity, one potential source that is not accounted for is difference in patentable subject matters across jurisdictions. Such differences would lead to a legitimate grant at one office and a legitimate refusal at another office, but would be interpreted as an error in one office. Whereas this point is certainly valid in theory, it is unlikely that applicants would file patent applications in jurisdictions where the subject matter is not patentable.

However, in an attempt to test the sensitivity of our results to differences in patentable subject matter, we report estimates by technology field. We know from discussions with patent examiners that patentable subject matters in mechanical engineering are very similar across jurisdictions, and this field will thus serve as our benchmark. In Table 7, we assign each family to one or more major technology OST technology groups based on any one of the IPC subclasses given at any office.[10] In addition, we use the 'Biotechnology' and 'Software' classifications from the OECD (2003) and Graham and Mowery (2004) respectively. Table 7 reports the predicted invalidity 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.

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

Table 7. Predicted invalidity rates, model M5 by technology fields

	EPO		USPTO		KIPO		JPO		SIPO	
	T	I	T	I	T	I	T	I	T	I
Electrical	5.2	5.3	15.7	4.2	12.8	4.2	2.6	5.8	15.6	1.8
Instruments	9.8	4.4	16.7	4.0	8.8	5.1	2.7	5.6	15.6	1.7
Chemicals & pharma.	16.3	3.1	14.8	6.3	9.4	6.0	4.6	7.2	19.7	1.6
Process engineering	11.2	3.5	14.2	5.0	9.2	5.4	2.6	6.1	17.9	1.4
Mechanical engineering	9.9	3.0	14.7	3.3	6.3	5.4	0.6	4.9	13.4	1.2
Biotechnology†	18.4	3.7	19.0	7.3	17.0	6.3	11.3	7.9	22.4	2.6
Software††	3.1	6.4	12.5	6.3	14.6	4.7	4.1	7.7	16.3	2.2

Notes: 'T' indicates differences that are due to differences in office threshold. 'I' indicates inconsistent decision at the focal office (weak patents). An application is allocated to one or more major technology groups from any of the 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 JPO is the office with the highest standards and the USPTO the lowest. Invalidity rates are 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 predicted invalidity rates seem misplaced. Second, if one believes that there is no unobserved heterogeneity within technology fields, then the estimates can be taken as reflecting differences in inventive step and invalidity rates across fields. The EPO applies particularly strict standards in software, and JPO is the strictest office in all other technology fields. The invalidity rate for biotechnology is higher than the base rate at all offices except the EPO, and the invalidity rate for software is higher than the base rate at all offices except the KIPO.

6.2 Out-of-sample validity

Patents applications in our sample are considerably less selected than in litigation studies previously used to study invalidity. In addition, compared to previous studies, the sample does not select on likely (in)validity or on invention quality. Our sample does 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 possibility that economic value may be correlated with inventive step 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 inventive step) 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 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 per cent 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 patents with higher than average inventive step.

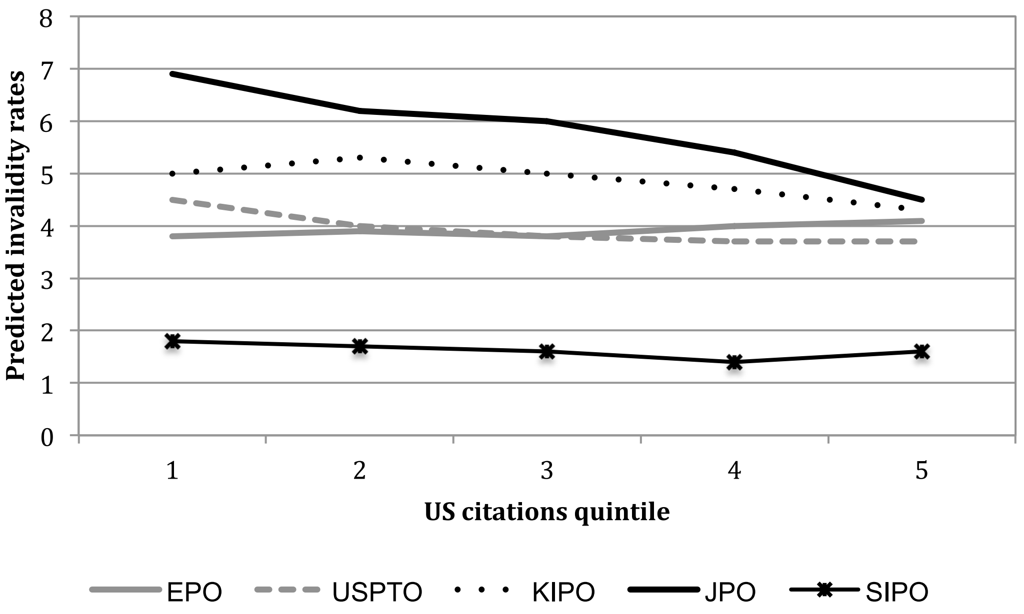
Table 8. Invention quality by filing status

	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 patents in our sample are somewhat selected on the quality of the underlying invention does not tell us anything directly about possible bias in our estimates of patent invalidity. We assess the effect of quality on invalidity 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 invalidity rates by quintiles of citation 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 invalidity rates go down from 7 per cent to less than 5 per cent.

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 invalidity rate is stable across quality. Therefore, the population-wide rate of weak patents is likely to be around 4 per cent. A similar reasoning holds for KIPO, with a population-wide rate of weak patents of about 5 per cent. There is strong selection into filing at SIPO but the invalidity rate is fairly stable across quality such that the population wide invalidity rate is probably close to 2 per cent anyway. In light of the strong selection into the filing decision at the USPTO, the population wide invalidity rate is probably closer to 5 per cent than 4 per cent. At the JPO, population-wide invalidity rate are probably closer to 7 per cent than 5 per cent for similar reasons.

6.3 Sensitivity to applicant experience

We next investigate whether the consistency of grant outcomes varies with the level of experience of applicants. On the one hand, more experienced applicants are presumably better equipped to push their patents through the examination process, leading to less mistakenly granted patents. On the other hand, more experience applicants may invest less energy in each patent, leading to potentially more heterogeneity in grant decision.

Figure 2. Accuracy of grant decision by applicant experience

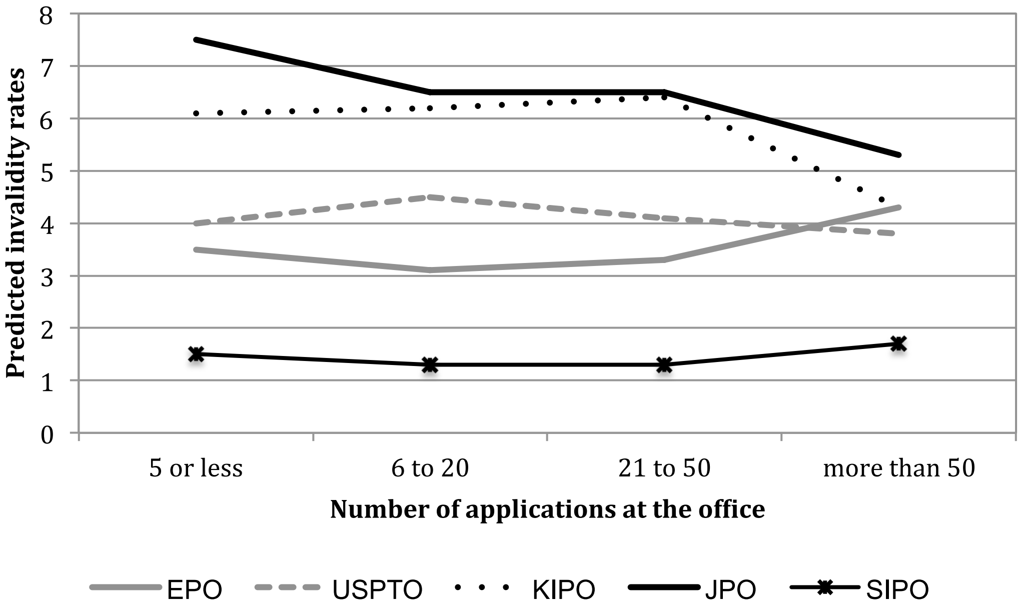


Figure 2 depicts the invalidity rates by applicant experience (measured in terms of the number of applications submitted to the focal office over the whole study period). Overall, no clear pattern emerges.

6.4 Additional considerations

We have also estimated model M5 on the subsample of 322,583 applications with the same number of claims across jurisdictions in an attempt to further control for unobserved heterogeneity. 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 took all Japanese applicants who filed at the JPO through the PCT but have no recorded legal status and recoded them as 'refused'. 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 is a quantitatively significant problem, with the overall fraction of dubious patents perhaps a quarter or more of all patents. Our analysis of patents 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 allows us to decompose differences in the decisions of multiple decision-makers into those that are due to an inconsistency or mistake by the first decision-maker, those that are due to a mistake by subsequent decision-makers, and those that are due to differences in the standard applied by different decision-makers. Note that the 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. 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 all three sources of inconsistent decisions are important.

We find that the fraction of Invalid patents—those that should not have been granted given the offices own grant threshold—does not exceed single digits for any office. While our sample is large, it is not randomly drawn. Patents examined in multiple international jurisdictions are likely to be of higher value than the average patent. Our analysis of the selection problem suggests, however, that invalidity rates for the population are unlikely to be much higher then our estimates for the sample. Thus, even allowing for selection bias, our results suggest invalidity rates much lower than the rates found by litigation studies. This suggests that litigated patents are highly selected towards those most likely to be found invalid or/and courts systematically apply a stricter standard for validity then the patent office. This is an important topic for further research, in order to clarify the implicit uncertainty about the likelihood that extant patents would survive a court challenge.

The fraction of patents that might be said to be low quality in the sense that they result from systematically low standards is larger, ranging from 9% for the EPO to approximately 11% for Korea and 15% for the United States and China. It is of course possible that all of these countries have standards that are too low, but commenting on that issue would require a normative analysis beyond our scope.

These findings 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. If the PPH agreements will increase the harmonization of decision, 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 other, which may create additional issues in the context of PPH agreements.

Our analysis is silent on the optimal level of ambiguity. 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 harmonised outcomes for businesses is likely to improve welfare. Our results provide a useful starting point in that regard.

Acknowledgments

The authors are grateful to seminar and conference participants at the Toulouse School of Economics and ETHZ for valuable feedback.

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.

Cremers, K., Ernicke, M. Gaessler, F. Harhoff, D. Helmers, C. McDonagh, L. Schliessler, P. Van Zeebroeck, N. (2013) Patent litigation in Europe, ZEW Discussion Papers, No. 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.

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, Michael and Wasserman, Melissa (2014). Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from MicroLevel Application Data, IP Statistics for Decisionmakers Conference, OECD and JPO, Tokyo

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. NBER Working Paper 21868.

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.

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.

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

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.

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.

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.

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

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