RIETI Discussion Paper Series 16-E-092

Contribution of Patent Examination to Making the Patent Scope Consistent with the Invention: Evidence from Japan

OKADA Yoshimi
Hitotsubashi University

NAITO Yusuke
Hitotsubashi University

NAGAOKA Sadao
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 16-E-092

November 2016

Contribution of Patent Examination to Making the Patent Scope Consistent with the Invention: Evidence from Japan[1]

OKADA Yoshimi (Hitotsubashi University)

NAITO Yusuke (Hitotsubashi University)

NAGAOKA Sadao (RIETI/Tokyo Keizai University)

Abstract

Delineating the patent scope consistently with the contribution of the disclosed invention is one of the most crucial requirements for a patent system to promote innovation effectively. Given the incentive of an applicant to set a scope for the patent as broad as possible, an important task of the Patent Office is to narrow it so that it becomes commensurate with the invention. This study analyzes empirically how significantly the Patent Office delivers this important function through patent examination, focusing on product patents in four major technology areas. We find that, often (i.e., two-thirds of the granted patents), the patent's scope is narrowed as an outcome of the patent examination. In addition, both the incidence and the extent of such narrowing increase when the applicant chooses broader claims and decrease when the quality of prior art disclosure by the applicant is higher, suggesting that patent examination indeed contributes to making the patent scope consistent with the contribution of the invention. We also found that a more important patent application experiences more the narrowing event, consistent with our simple model of examination where an examiner aims at reducing the economic cost due to excess claim such as deadweight loss, subject to time constraint.

Keywords: Patent scope, Patent examination, Quality of patent, Narrowing of patent scope, Claim length, First claim, Quality of prior art disclosure

JEL classifications: O34, O31, O38, K29

[1]This study is conducted as a part of the Project "Study on Industrial Innovation Capability and Innovation Infrastructure" undertaken at Research Institute of Economy, Trade and Industry (RIETI).

1. Introduction

Delineating the scope of a patent consistently with the contribution of the disclosed invention to the state of the art is one of the most crucial requirements for a patent system to promote innovation. A patent holder specifies the scope of his/her exclusive rights in "claims." An overbroad claim restricts the use of knowledge already in the public domain or preempts knowledge not yet invented, whereas a claim that is too narrow weakens the incentive for an invention that has large inventive step. It is important to note that applicants have an incentive to make their patent claims as broad as they can as broader claims increase the exclusionary power of patents and therefore their profit from the inventions. An important task of the Patent Office, therefore, is to narrow down the overbroad claims to ensure they are consistent with the invention. That is, a patent examination process is not only to assess the patentability of the invention, but also to guide the applicant so that they sufficiently and distinctly specify in the claim the difference between their invention and prior arts and to narrow down the claim so that it is commensurate with the contribution of the invention to the state of the art. This study analyzes empirically how the Patent Office serves this important function through its patent examination process.

This role of the Patent Office has been at the heart of recent concerns over low quality patents (FTC (2003); Jaffe and Lerner (2004); National Academies (2004); GAO (2013)). Probabilistic patents with uncertainty in their patent validity can result in a situation where a "weak" patent can obtain strong market power (Farrell and Shapiro (2008); Lemley and Shapiro (2005)). A large number of potentially invalid patents increase clearance costs, making inadvertent infringement more likely and litigation more frequent (Bessen and Meurer (2008)). Recently, governments in major countries have launched policies to increase patent quality in order to promote innovation. For example, in 2013, the Economic and Scientific Advisory Board (ESAB) of the European Patent Office (EPO) issued a statement with recommendations for improving the patent system, as well as a report that highlighted the importance of patent quality in boosting innovation. In Japan, the Cabinet decided in July, 2014, to aim to achieve the highest quality patent examination in the world through several measures, including introducing an objective quality management system. On February 4, 2015, the US Patent and Trademark Office (USPTO) launched “Enhanced Patent Quality Initiative,” which consists of comprehensive measures to improve patent examination. Thus, it has become one of the top priority tasks for these patent offices to ensure that the patents issued are not overly broad, that claims are legitimately novel and non-obvious, and that the scope of the patents is distinctly specified and clear.

To the best of our knowledge, however, there has been no full-scale econometric analysis of how the Patent Office actually contributes to making a patent scope consistent with the invention through applicant’s amending the claims. Most economic analyses of the patent system have not considered such amendment, assuming the task of a patent examiner to be simply to assess the novelty, the inventive step, and the utility of the invention, and to conclude whether the invention is patentable or not (see Schuett (2013) for an example). However, this misses a crucial function of the patent examination, which is to guide applicants in delineating the scope of the patent properly so that it is consistent with the invention through the amendment of its claims.

This study investigates to what extent the Patent Office contributes to making the scope of patent rights consistent with the contribution of the invention, based on a large scale patent examination data on product patents in four major technology areas. We use the claim length (or, more exactly, the inverse of the claim description length) of the first claim (claim no. 1), as well as the number of claims, as a measure of the breadth of the scope of a patent right. This study also develops a new measure of the quality of the disclosure of relevant prior art by the applicant, which we use as a proxy of the excess claim by the applicant. If the examination process is working properly for making the scope of patent rights consistent with the contribution of the inventions, the broader initial claims will be narrowed more frequently and to a greater extent. In addition, the patent applications with higher quality disclosure of prior art are less likely to undergo narrowing down processes and, if they do, the extent of narrowing will be smaller, because such applications are more likely to have claims consistent with the contribution of the disclosed invention to the state of the art. Furthermore, patent applications with higher value or importance are more likely to undergo narrowing processes because patent examiners are likely to devote more effort to reduce the excess claim of a patent that could result in a large economic cost.

In summary, we found that these hypotheses are significantly supported. The scope of patents are often (around two-thirds of all patents granted) narrowed as an outcome of the patent examination process. An applicant is more likely to be granted a patent with the initial claim intact if the initial claim is narrower and the prior art is well disclosed in the patent application; and, if such a claim is amended, the extent of narrowing is likely to be small. Furthermore, an important patent is more likely to be narrowed.

This paper is organized as follows: section 2 presents a short description of the legal background and prior literature; section 3 provides a simple model of patent examination and derives three hypotheses; section 4 offers explanations of our novel comprehensive database of the Japanese patent applications and grants, its basic statistics, and the estimation models for the incidence of the narrowing of the first claim and for the extent of narrowing, based on instrumental variables; and section 5 presents estimation results and discussions. Section 6 then concludes the study, with a discussion of policy implications.

2. Legal background and prior empirical literature on patent scope

The scope of a patent right is described in documents called “claims.” The process of a patent examination up until the grant decision can be summarized as follows. First, an applicant makes an initial claim. Next, a patent examiner presents prior art that denies the novelty or inventive step of the focal claim. The applicant then amends the claim to narrow down the scope of the exclusive right so as not to cover the prior art. Finally, the patent examiner grants a patent if the amendment is appropriate.

There are two typical ways to narrow down the scope of a patent[2]. The first is to add new elements to restrict the scope of the concepts of the claim. This is the most frequently used way for inventions of appliances. The scope of the invention is presented by the logical multiplication of the scope of each item. In this case, each added word has a role of limiting the scope of the patent right. We call this type of claim structure the "cascaded type." The other way is to eliminate some of the items that are listed in the claim as the choices that a person will need to make when using the invention; this is frequently used for inventions of materials or chemical compositions, especially in inventions described by Markush-type claims. The scope of the invention is presented by the logical addition of the scope of each item. We call this type of claim structure the "parallel type."

Typically, the first claim conveys the broadest inventive concept and so this is the one we focus on. Because the Japanese language does not use spaces between words, it is difficult to count words automatically in Japanese sentences; we therefore use the number of characters instead of the number of words to specify the length of a claim. In this paper, we define "claim length" as "the number of characters in the first claim of a patent application or grant."

In an independent cascaded-type claim, the number of elements that limit the scope of an inventive concept increases with the number of characters; thus, in this type of claim, there is a negative correlation between the number of characters and the breadth of the conceptual scope. Accordingly, we measure the breadth of the claim by the inverse of the claim length.

There have been only a limited number of studies that have investigated examination

[2] There is another possible way to narrow down a claim, which is to replace a general term by a more specific term. However, it is quite rare that a patent can be obtained only through this type of an amendment. We have therefore ignored this type of amendment in this study. The three ways are not mutually exclusive and an applicant can make an amendment that involves a combination of these narrowing processes.

outcome, focusing on the patent scope. Malackowski and Barney (2008) used the average number of words per independent claim in a US-issued patent as a proxy for the quality of patent examination by the USPTO. They noted that the number of words in independent claims increased as a result of examinations; around the time of their investigation, the mean number of words increased from 111.1 on filing to 153.2 on issuance. They also found that independent claims for US patents issued in 2007 had 4.4 percent more words than those issued in 2003, which they considered to be evidence that the quality of patent examinations had not decreased during that period. Although they used the claim length as a proxy for the quality of the patent examination procedure, they did not identify the factors causing the increase in the number of claim words during the patent examination.

Harhoff (2016) used the inverse of the length of the first independent claims as a proxy measure for patent breadth. He pointed out that the inverse of the patent breadth of European Patents has become significantly longer over the years: 165 words in 1990, 175 words in 2008 and 183 words in 2014. He attributed this trend toward narrower scopes to the reform of the fee structure and the restriction of divisional applications. He did not investigate the factors that this paper focuses on.

Abiko and Tanaka (2010) investigated how the claim length changed between the initial claim for a patent application and the claim at the time of grant in the technological field of optical communication (International patent classification: H04B). They found in this technological field that all the mean values of claim length of granted patents calculated by each applicants of the top 10 patent holders in this field are greater than those of applications at filing. However, they did not investigate the causes. Webster et al. (2014) investigated the incidence of change in independent claims in 236 European patents and 82 Japanese patents. They found that foreign inventors filing patent applications at the EPO and Japan Patent Office (JPO) were more likely to receive a grant with claim changes than were domestic inventors. However, the focus of their investigation was the violation of the national treatment principle stipulated in the Agreement on Trade-Related Aspects of Intellectual Property

Rights. Conversely, Osenga (2011) focused on the number of words in a claim and used it as a proxy for the readability or comprehension of claims.

Regarding applicants' disclosure of prior art, Cotropia et al. (2013) investigated the references of US patents that examiners used in their office actions to narrow patents. They found that only 12.7% came from the applicants, with the remaining 87.2% coming from the examiners, suggesting that the US doctrine of the strong presumption of validity enjoyed by patents during litigation should be modified to reflect the realities of patent examination. Sampat (2010) investigated US patents and their citations and concluded that applicants search to a greater extent for prior arts and disclose more prior arts for their important patents. However, these two studies did not investigate the effect of the applicants' prior art disclosures on the outcome of the examinations, in particular, on the narrowing of the patent scope.

Lemley and Shapiro (2005) suggested that devoting substantially more resources to patent examination is more likely to be efficient if the additional resources can be focused on patents whose validity will turn out to be commercially significant, and that patent examiners might focus greater attention on the patent applications with features that correlate with greater ultimate value. However, as far as we know, there has been no empirical study that has provided such evidence.

3. A Model of patent examination and hypotheses

To construct a simple economic model of the patent examination, we consider the following simplified examination process. An applicant files a patent application. An examiner then assesses whether the claim of the patent is consistent with the contribution of the disclosed invention to the state of the art. If the claim is too broad relative to the net contribution of the invention, the examiner issues an initial office action of rejection, which forces the applicant to amend the claim so that it does not cover what is already in the public domain. At the same time, the examiner needs to spend time resources on finding the appropriate prior art and persuading the applicant, given that the burden of proof for narrowing the claim resides with the examiner.

Given this institutional setting, we consider the following simple model of the patent examination. We take an applicant's behavior as given, parameterized by the scope of the initial claim y_{int} and the extent of excess claim (θ).We denote the ideal scope, i.e., the broadest legitimate scope when all the prior art has been found and identified appropriately, by $y^* = y_{int}(1-\theta)$, where θ represents the proportional extent of the excess claim by the applicant relative to the ideal scope. Such a gap may arise because the applicant does not have to search the prior art before the application whereas the examiner is an expert in prior art searches and has to search the prior art comprehensively. In addition, the examiner can exploit not only the applicant's disclosed prior art but also all other knowledge sources on the trajectory of the technological development up to the date of the application. The gap may also arise when an applicant strategically specifies a broader claim (Lampe (2012)). We introduce y, which represents the potential patent scope after the amendment.

An excess claim amounting to $(y - y^*)$ would allow a patentee to exercise excess additional market power with regard to the patented technology. Such market power can cause a deadweight loss, an unjustified loss of consumer surplus, and possibly an unproductive rent dissipation loss such as the cost of a competitors' investment to invalidate the patent or to invent around the excess claim. We assume that the examiner would like to reduce such economic losses. An individual examiner's incentive for this may come from a sense of responsibility as a public servant, from the satisfaction obtained by fulfilling the role expected by Patent Law, or from the intrinsic motivation to undertake a high-quality examination and search of prior art. The examiner considers both the social benefit of narrowing the initial claim y_{int} toward y^* and the cost of spending additional time resources to achieve this. The marginal cost (MC) incurred by an examiner (the Patent Office) for the additional effort in search and persuasion (that is, the marginal cost for narrowing) increases as the amended claim y approaches y^*. Furthermore, when the amended claim y is sufficiently close to y^*, the *MC* becomes greater than the marginal benefit (*MB*) of additional reduction of the excess claim, because the latter is close to zero, as explained below in the case of deadweight loss, while identifying prior art becomes more difficult as y approaches y^*.

We assume for simplicity that the excess claim amounting to $(y - y^*)$ allows the applicant to increase the price p of its patented technology by $\alpha(y - y^*)$ above the price ($p^* = \alpha y^* = p(y^*)$) that would have prevailed had there been no such excess claim, where α translates the patent scope into the price of the technology or the consumers' willingness to pay[3]. This price-increasing effect exists because the excess claim covering the prior art makes it difficult for competitors to sell substitutes closer to the patented technology. The excess price increases the deadweight loss of the patent by reducing the use of the technology, as follows. If we assume that the demand for the technology is given by $Q = sq(p)$, where s represents the size of the market (the number of potential customers) and $q(p)$ represents the proportion of the customers who actually use this technology for price p. Then, the marginal increase of the patent scope by $\mathrm{d}y$ results in a price increase of $\alpha \mathrm{d}y$, which in turn results in a decrease in the demand for the use of the technology by $\mathrm{d}Q = s\partial q/\partial p \alpha \mathrm{d}y < 0$. The increase in the deadweight loss is given by $p\mathrm{d}Q = \{\alpha(\mathrm{y} - y^*) + p^*\}dQ$ (the loss is negative), assuming that the marginal cost of using the technology is 0. Because the Patent Office would be concerned only with the part of the deadweight loss due to the excess claim[4], the *MB* for reducing the excess claim can be given by

$$MB = \alpha(y - y^*)dQ = \alpha^2(y - y^*)s\partial q/\partial p dy \qquad (1)$$

Denoting $v = \alpha^2 s$ as a measure of the economic importance of the patent scope[5] and

[3] Note that inventions with the same claim length can vary significantly in technological or economic importance.
[4] Otherwise, the Patent Office would be inclined to reduce the patent scope further below y^*.
[5] According to Okada et al. (2016), the inverse of the number of characters in the first claim has a significant explanatory power for the values of patents measured by the

further assuming that the price sensitivity of demand is constant (assumed to be $-\beta$), the *MB* of narrowing for the examiner is given by

$$MB(v,\ (y-y^*)) = \beta v \cdot (y-y^*) \qquad (2)$$

The *MC* of narrowing for a patent examiner, on the other hand, increases as the amended claim y approaches y^*, as given by the following function of $(y-y^*)$:

$$MC(y-y^*) = g(y-y^*) \qquad (3)$$

with $MC' = g' < 0$. The cost to the examiner can be considered to be the loss of the amount of time that would be used for examining other patent applications. Here we assume that the marginal cost does not depend on the economic importance of the patent (the implication of relaxing this assumption will be discussed later in section 5 on estimation results and discussions).

Then, the optimal patent scope the examiner would choose is given by

$$MC(y-y^*) = \beta v \cdot (y-y^*),\ \text{where}\ y^* \leq \text{y} \leq y_{int} \qquad (4)$$

given that the second order condition is satisfied (that is, the marginal cost curve for narrowing is downward sloping while the marginal benefit of narrowing is upward sloping with respect to the patent scope; Figure 1). If the marginal cost of examination is zero (that is, $MC(y-y^*) = 0$), the optimal scope becomes y^* regardless of the importance of the patent.

Figure 1 here

Equation (4) gives the threshold of the scope of the initial claim below which the application would not invite the examiner's rejection:

$$y_{threshold} - y^* = f(v)\ \text{with}\ f' < 0 \qquad (5)$$

This threshold declines with the importance of the invention v, given the second order condition. If the initial claim exceeds this threshold, an office action of rejection will be made to lead the applicant's

---

number of applicant forward citations.

amendment amounting to the following, given that the applicant wishes to have the largest admissible scope of the claim (unless the applicant abandons the application):

$$y_{int} - y_{threshold} = y_{int} - y^{*} - f(v) = y_{int}\theta - f(v) > 0 \text{ with } f' < 0 \qquad (6)$$

If we further assume, for simplicity, that $MC(y - y^{*}) = k/(y - y^{*})$, we then have the following explicit solution for the extent of the narrowing when it takes place:

$$y_{\text{int}} - y_{threshold} = y_{\text{int}} - y^{*} - f(v) = y_{int}\theta - \sqrt{k/(\beta v)} > 0 \qquad (7)$$

We have explained the case in which the examiner aims at reducing an excess deadweight loss, but the conclusion does not change even if we use an unjustified loss of consumer surplus or an unproductive rent dissipation loss as a measure of the economic cost, provided the *MC* for the examiner of narrowing the claim becomes very large toward y^{*} and crosses the *MB* curve from below as *y* declines. Note in this regard that these two economic costs due to the excess claim also increase with the importance of the invention (the size of the market s and the price effect of patent scope).

Thus, we have the following proposition on examination, which is directly implied by equation (6) as well as Figure 1.

Proposition on the examination outcome

The claim is more likely to be narrowed if the initial scope of the claim (y_{int}) is large, the extent of the excess claim (θ) is large, or the importance of the patent application is large (large v). Similarly, the extent of narrowing increases with the initial scope of the claim (y_{int}), the extent of the excess claim (θ), and the importance of the patent scope (v).

Up to this point, we have focused on the inverse of the description length of the first claim as a measure of the scope of a patent right. A further proxy, the number of claims, can measure the scope of the right. When the first claim is broader, there will be more embodiments of the claimed concept and there will therefore tend to be a greater number of claims. At the same time, for a given scope of the first claim, a patent with a greater number of claims has a broader patent scope.

Although y^* is unobservable, unlike y_{int}, we use the quality of the applicant's disclosure of prior art as a proxy for θ, which indirectly provides us with y^*. This is because an application with more disclosures of relevant prior art (i.e., a higher-quality disclosure) is more likely to have claims consistent with the contribution of the disclosed invention to prior art (i.e., a lower θ). As a measure of this quality, we introduce q, the share of co-cited patent literature (that is, prior patent literature cited by both the applicant and the examiner) out of all the patent literature citations made by an examiner that examined the focal patent, which we will explain later in detail, including the issue of addressing the endogeneity of this variable.

Given this empirical implementation, we can formulate the following three hypotheses, which we will examine empirically.

Hypothesis 1. The effect of the breadth of initial claim on the examination outcome

A patent application with a broader initial claim (either with a shorter claim length or with a larger number of claims) is more likely to undergo narrowing. In addition, if an amendment is made, the extent of narrowing is greater when the initial claim is broader.

Hypothesis 2. The effect of the quality of the applicant's disclosure of prior art on the examination outcome

A patent application where the applicant's disclosure of relevant prior art is of higher quality is less likely to undergo narrowing. In addition, if an amendment is made, the extent of narrowing is less for an application with higher quality disclosure of prior art.

Hypothesis 3. The effect of the importance of a patent on the examination outcome

A patent application of greater importance is more likely to undergo narrowing. In addition, if an amendment is made, the extent of narrowing is greater for such an application.

4. Data construction, descriptive statistics, and estimation models

4.1. Data construction

Our Japanese patent database contains the number of characters in the first claim, as well as bibliographic and examination data. The data were newly developed from the two Japanese patent databases maintained by Artificial Life Laboratory, Inc. The first of these, “td-5,” comprised the text data of the pre-grant patent application publications and the granted patent publications; the second, “Pat-R,” comprised the bibliographic and examination process data. Both were originally published by the JPO.

The length of the first claim may measure patent scope differently between product and process patents. This paper focuses on product patents, which account for more than 80% of patents (as will be explained in the following subsection). We automatically divided the inventions of the first claims into the two categories by analyzing the preamble and the final part of the first claims, where the statement relating to the category of the invention usually exists in the Japanese language. We checked the results and amended the program several times. After this adjustment, we randomly chose 460 sample patents and verified that there were no errors in the machine-identified category data of the first claims[6].

Applicant citations are documents disclosed by the applicants in the specifications of the

---

[6] We removed from our data several patents with fewer than 10 characters in the first claim, which included first claims indicated as “removed.”

patent applications. Applicant citation data, obtained by searching the full specifications, were included in the td-5 database. To address the variations in the notation of applicants' names, we used the National Institute of Science and Technology Policy (NISTEP) dictionary of Japanese company names and the connection table to the Institute of Intellectual Property (IIP) patent database, both of which were provided by NISTEP. Each major applicant has a unique NISTEP identification (id). By using the id data in the NISTEP dictionary, we were able to identify the company even if there were changes in the company name, or mergers and acquisitions. We only considered applications with applicants identified by the NISTEP applicant id, i.e., only Japanese companies. Co-owned patents were removed in order to use a fixed effect applicant model in regression analyses.

We used the European Patent Office Worldwide Patent Statistical Database (PATSTAT) (2014, autumn version) to obtain information about the existence of patent families, namely the corresponding US patents that were filed at the USPTO and granted patents for the same inventions. We used the Japanese patent applications only for those with filing dates between January 1994 and August 2002, dates chosen to eliminate the influence of a legislative change on the amendment of claims in January 1994 and a legislative change on disclosure in September 2002. For simplicity, we also eliminated divisional applications from our datasets.

We divided technological fields according to the IIP patent database, which divides technological fields into 33 categories following the WIPO classifications (See Goto and Motohashi (2007)). We then grouped the 33 technological fields into six large technology fields, namely, "Chemical" (Chem.), "Computers & Communications" (C & C), "Drugs and Medical" (D & M), "Electrical & Electronic" (E & E), "Mechanical" (Mech.), and "Others" (see Appendix I for the International Patent Classification (IPC) correspondence). This paper focuses on four sectors, C & C, E & E, Mech. and Others, where cascaded-type claims are expected to be dominant and the amount of parallel type claims is relatively small. These four sectors account for 86 % of the entire granted patents of our sample.

We also eliminated patents and patent applications that involved chemical, mathematical formulas, or tables in the first claim, utilizing the tag and related information in the patent database; the crucial part of the patent claim is provided by image data rather than text data in these patents[7].

4.2. The distribution of claim length of applied and granted patents

Table 1 presents descriptive statistics for the claim length on granted patents in each technological field. Product inventions accounted for more than 80% of the inventions in each of the four fields. C & C included the patent grants with the longest average claim length. Furthermore, the mean claim length was longer for product inventions than for process inventions in all major fields. The composition between product inventions and process inventions also differed between the technological fields. Hereafter this paper focuses on product inventions. Figure 2 shows the distribution of the natural logarithm of the inverse of claim length of granted patents for product inventions in all fields (the aggregation of product patents in C & C, E & E, Mech., and Others). The shape of the distribution curve is close to that of a normal distribution.

Table 1 here.

Figure 2 here.

Figure 3 shows a comparison of the distributions of the natural logarithm of the inverse of the claim length before and after examination for product inventions in the C & C and E & E sectors. The solid lines with diamonds indicate the initial claim lengths at filing and the dashed gray lines with triangles indicate the initial claim lengths at filing for the applications that were eventually granted; the lines with circles indicate the claim lengths of granted patents. It is evident that all the distribution

[7] Less than 5% of the data in each field was eliminated, with the following rates for the product patent data: overall, 1.6%; C & C, 0.4%; E & E, 3.6%; Mech., 1.0%; and Others, 1.1%.

curves are approximately normal in distribution and that the curves move significantly to the left after examination, indicating that the average scope of the granted claim after examination became narrower than that of the initial claim.

Figure 3 here.

4.3. Examination outcomes by the initial claim length

Tables 2A and 2B show the outcomes of patent examinations for the four sectors. "Not granted" includes rejected, withdrawn, abandoned, or dismissed after the first actions of patent examiners[8]. Granted cases are divided into three categories–"broadened," unchanged, and narrowed–according to the change in the scope of the claim between the initial first claim and the first claim at grant, corresponding to cases where the claim length decreased, remained unchanged, or increased, respectively. The claim length decreased in only a minimal number of cases (around 3 to 5 % of the total granted patents), suggesting that we effectively succeeded in eliminating the patents with parallel-type claims from our sample. However, there remains a significant possibility that this was incomplete and that the decrease in claim length was still an outcome of narrowing down the scope by eliminating choices or alternatives written in an initial parallel-type claim[9]; we have therefore removed those patent applications from our sample for the following estimations. In the majority of cases, the claim length increased after the patent examinations (68% of the granted patents in the examined sectors), with the ratio varying only slightly across the four sectors. Thus, it is evident that patent examination plays a very significant role in limiting the scope of the patent in all these technology fields.

Tables 2A and 2B here.

Figures 4A and 4B show how these examination outcomes differ by initial claim length for

[8] Even after a decision is made to grant a patent, the application may still be dismissed if the applicant does not pay the registration fee.
[9] We bracketed term ("broadened") rather than broadened to signal this possibility. On the other case, there are cases where the applicants broaden the initial claims after finding that the prior art are scare even considering those found by the examiners, although they seem to be very rare.

the C & C and E & E sectors. The figures show that the grant probability decreases and the narrowing probability increases significantly with the breadth of the scope of the claim (i.e., the inverse of claim length) at patent application. This indicates, as expected, that the patent examiner is more likely to narrow the claim when the applicant pursues a broad claim (a claim with a short description). We will show later by regression analysis that such narrowing is more likely to occur when the initial claim is broad (either a first claim with a short claim length or a large number of claims), controlling for the quality of applicant's disclosure of relevant prior art (the proxy for θ) and proxies for the importance of the filed invention (v).

Figures 4A and 4B here.

4.4. Disclosure quality and estimation issues

Before we use examiner citations as the basis for evaluating the quality of applicant's disclosure of relevant prior art, we shall first briefly explain them. Examiner citations are made when a patent examiner makes office actions. There are three major types of office actions. The first is "notification of reasons for refusal," in which a patent examiner cites prior art documents as a basis for rejecting claimed invention(s) on the ground that they lacks novelty or an inventive step. A patent examiner may also cite prior art documents that do not directly constitute a basis for rejection in this kind of office action in order to give the applicant the opportunity to make an effective amendment. The applicant can make an amendment after receiving notification of the reasons for refusal. A patent examiner is expected to cite the closest prior art or the closest set of prior arts in notification of the reasons for refusal. The second type of office action is "final rejection," and a patent examiner may cite additional prior art to reinforce the basis for rejection. The final type of office action is "decision to grant." In this case, the patent examiner may cite prior art as references if he/she considers that it is beneficial for the public that these references are indicated in the publication of the granted patent. In the great majority of cases, one or more notifications of the reasons for refusal are made before the decision to grant.

Figure 5 shows our measure of the quality of applicant's disclosure of prior art, *q*, which we have developed by matching examiner and applicant citations. In the figure, the citation category *a* is the number of prior arts cited only by the applicant, *b* is the number of prior arts cited both by the applicant and by the examiner, and *c* is the number of prior arts cited only by the examiner. If we define $q = b / (b + c)$, *q* can be interpreted as the quality of the applicant's disclosure of prior art, as measured by the coverage of examiner citations. That is, if the applicant anticipates and discloses all the examiner citations, the disclosure quality is 1 (100%). Conversely, if the applicant anticipates none of them, it is 0, even if the application refers to prior art[10].

Figure 5 here.

This measure of the quality of prior art disclosure leads to two estimation problems. First, the measure is likely to be endogenous, given that it is also affected by the decision made by the patent examiner who examined the focal patent application, i.e., by how effectively the examiner identified the relevant prior art for the focal patent application. If an examiner happens to fail in finding some relevant prior art, the variable *q* is high not because of the high quality of the applicant's disclosure, but due to *c* being small due to the examiner's mistake. At the same time, such an application is less likely to undergo narrowing; thus, we would observe a negative correlation between variable *q* and the narrowing. This would be a spurious correlation. Such endogeneity would make the ordinary least squares (OLS) regression overestimate the effect of the quality of prior art disclosure.

Second, *q* is an imperfect measure of the quality of prior art disclosure because it is based only on the share of the co-cited patent literature. In particular, it neglects the content of prior art and

---

[10] We consider only Japanese patent literature due to non-availability of the equivalent data for non-Japanese patent literature in the database.

hence the variations in prior art in terms of its importance in determining the contribution of the invention. In addition, there can be those cases in which the examiner found other prior art documents that are as relevant as those disclosed by the applicant, and he/she has full discretion in citing which of them. In such cases, q may .differ even if the quality of applicant's disclosure is equal among such cases. If such measurement errors are significant, the OLS regression underestimates the effects of the quality of the applicant' prior art disclosure.

We introduce two instrumental variables to address these two problems: the past (one-year average) quality of prior art disclosure, and the past (one-year average) number of applicant citations. Both are calculated for the period of the preceding twelve months for each combination of applicant and technology field pair (for the six broad technology fields). We eliminate the patents of the applicants who obtained fewer than 250 patents in the focal sectors with a filing date between January 1993 and August 2002 to implement instrumental variables (IV) estimations. Given that each patent is examined independently of past patent examinations (note that we removed divisional applications from our sample) and that an examiner is randomly assigned to a new patent application in a given technical field that is much narrower than the six broad technology fields, there should be no correlation between the past average disclosure quality of each applicant in terms of q and the examiner's performance in examining the current patent application. However, there is likely to be a significant correlation between the two instruments with the quality of applicant's disclosure of prior art in the current patent application, reflecting the persistence of the disclosure quality of the applicant in each technology field.

4.5. Estimation models

The first estimation model is a model of the narrowing event, as follows:

$\text{Dummy for the } incidence\ of\ narrowing\ = \beta_0 q$

$$+ \beta_1 \, ln\_invers\_claim\_length(initial) + \beta_2 \, ln\_N\_claims(initial)$$

$$+ \beta_3 \, ln\_N\_inventor + \beta_4 \, ln\_N\_Forward\_Citation + \beta_5 \, US\_Grant\_dummy$$

$$+ \beta_6 \, subAppln\_dummy$$

$$+ effective\_filing\_year\_dummies + technology\_dummies$$

+ *effective_filing_year_dummies* × *technology_dummies* + (applicants' FE) + const + ε *(8)*

The dependent variable is the dummy for the narrowing event, which is set to 1 if the claim lengthens after the examination and to 0 if the claim length does not change. Among the explanatory variables, variable *q* is the coverage of the examiner citations by the applicant citations. As instruments, we use the past one-year average of *q* and variable *ln_av_N_Backward_Citation*, which represents the natural logarithm of the past one-year average of the number of applicants' backward citations for each technology sector. We utilize the following two variables to measure the breadth of the initial claimed scope: *ln_inverse_claim_length(initial)* and *ln_N_claims(initial)*. The former stands for the natural logarithm of the inverse of the initial (i.e., at the time of patent application to the JPO) claim length (the number of characters in the first independent claim). The latter stands for the natural logarithm of the initial number of claims. We utilize the following variables to measure the technological or economic importance of the inventions: *ln_N_inventor*, the natural logarithm of the number of inventors; *ln_N_Forward_Citation*, the natural logarithm of the number of applicants' forward citations; *US_Grant_dummy*, which is set to 1 if there is at least one corresponding US patent and to 0 otherwise; *subAppln_dummy*, which is set to 1 when there exist subsequent divisional applications for the focal patent and to 0 otherwise. We use *effective_filing_year_dummies* to represent the earliest priority year and technology dummies that correspond to the 33 technology classifications and their interaction terms. We use the fixed effect model of applicants, controlling for company-level missing variables[11]. Descriptive statistics for the data are presented in Table 3.

[11] One potential missing variable is the capability of the intellectual property management of the

Table 3 here.

The second estimation model is a model for the extent of narrowing. The dependent variable is a continuous variable, *extent_of_narrowing*, which indicates the magnitude of change in the inverse of claim length due to the examination, defined by the following formula:

extent_of_narrowing = ln(inverse of initial *claim length*–inverse of *claim length* at grant) (9)

This estimation model uses all the explanatory variables used in the previous model with the exception of *ln_inverse_claim_length*, as the dependent variable *extent_of_narrowing* directly uses this variable and so would tend to generate a negative correlation between the claim length at application and the measure for the extent of narrowing[12]. The sample for estimating equation (9) consists of the granted patent that underwent amendments to narrow down the first claims.

5. Estimation results and discussions

We present OLS and IV results for the *incidence of narrowing,* and then for the *extent_of_narrowing* and robustness check. All estimations were implemented using the fixed effect model of applicants. There are two IV methods. The first uses both the average of past *q* calculated for the preceding twelve months and the logarithm of the average of the number of the applicant's backward citations for the same period, each for all dyadic combinations of an applicant and a technology field (out of the six broad technological fields). The second IV method uses only the first instrument.

5.1. The incidence of narrowing events

---

applicant company. A company with a high score for such a capability would apply for patents with broad claims and could simultaneously have a small probability of amendment for narrowing the claim.

[12] We implemented the regression with adding *ln_inverse_claim_length* as an independent variable to the model and verified that the sign and the significance of the estimated coefficients of *q* and *ln_N_claims(initial)* do not differ significantly from those of the regression results based on the model without *ln_inverse_claim_length* as an independent variable.

Table 4 shows the results of the OLS and the IV estimations of the incidence of narrowing against *q*, the two claim variables at application, the variables to measure the importance of the inventions, and the control variables including the applicant dummies (the applicant fixed effects). The estimation is based on data pooled over four technology sectors. The results of the regression of the first stage (estimates (3) and (5)) indicate that the instrumental variables significantly explain the variation of variable *q*, showing that the disclosure policy of an applicant has significant persistence over time.

Table 4 here.

The coefficients of *ln_inverse_claim_length* and *ln_N_claims* are both positive and statistically significant at the 1% level for all three estimations (the OLS and the two IV estimations). Furthermore, the estimated coefficients are similar for the three estimates. Thus, it is evident that a patent application with a broad claim (in terms of either the inverse of the description length of the first claim or the number of claims) is significantly more likely to undergo a narrowing decision by the examiner, supporting Hypothesis 1. The elasticity of the claim length is large (around 0.2): a one standard deviation reduction in the number of characters of the first claim (around 0.5 in logarithmic terms) will result in an increase of around 8–10 percentage points in the probability of a narrowing event. For the number of claims, a one standard deviation increase (0.8 in logarithmic terms) will result in an increase of 3-4 percentage points in the probability of a narrowing event.

The coefficient for the quality of prior art disclosure, *q*, is negative and statistically significant at the 1% level for all estimates, suggesting that when an applicant cites more prior patent literature to be cited also by the patent examiner, it is more likely that the claim length will remain unchanged. Higher quality disclosure, which is likely to be accompanied by claim drafting that is more consistent with the contribution of the invention to the state of the art, reduces the probability of a narrowing event for the claim, supporting Hypothesis 2. The size of the estimated coefficient is much larger for the estimates by IV estimations (more than 5 times larger): -0.13 for the OLS estimation (estimate (1)) vs. -0.70 and -0.80 for the IV estimations (estimates (2) and (4), respectively). According to IV estimate (3), a one standard deviation (0.16) increase in the quality of prior art disclosure results in a decrease of 11 percentage points in the probability of a narrowing event, showing that the effect of the applicant disclosure on the incidence of the narrowing event is large. The significant underestimation by the OLS also indicates that there exists a serious measurement error in our measure of the quality of the applicant's disclosure of relevant prior art, which tends to dominate the endogeneity bias for estimating *q*.

The forward citation has a positive and highly statistically significant coefficient of around 0.02. There are no significant differences between the OLS and the two IV results for the estimated coefficient. The number of inventors and the dummy for the existence of a US corresponding patent also have positive and highly statistically significant coefficients (0.005 to 0.007 and 0.005 to 0.009 for the OLS and two IV estimations, respectively). These results imply that a patent application of greater importance is more likely to face initial rejections for amendments, after controlling for the other variables, supporting Hypothesis 3. In our economic model of patent examination, we assumed that the marginal cost of narrowing the scope of a patent is independent of the importance of the patent scope. It may rise, however, with the importance of the patent scope for the following reasons: examining an important invention may need searching divergent fields, and is thus more costly; or the search itself is an exploratory process involving trials and errors because no technological fields have yet been established, and it is thus costly. A rise in MC due to the increase of the importance of the invention will cause an increase in the threshold, which results in a reduction of the incidence and extent of narrowing. Our results are opposite to this prediction, which means that the importance of the patent scope affects the *MB* curve significantly more than the *MC* curve, and examiners consider the importance of inventions highly significantly in deciding how much effort they exert for their examinations.

The coefficient of *subAppln_dummy* is positive and statistically significant at the 1% level. The value is high and close to 0.1 for the OLS and IV estimates. This finding indicates that a patent with a future divisional application has a more than 10% higher probability of undergoing a narrowing event. Under the Japanese patent law system before April 2007, an applicant was not allowed to file a divisional application based on the focal patent after he/she received the decision to grant a patent; the patent law was subsequently revised to allow this. Therefore, if an applicant wants to reserve the opportunity of filing a divisional application based on the focal patent, he/she makes the initial claim broader so that the first action made by the patent examiner is notification of the reasons for refusal. There is another explanation. An applicant sometimes acquires a focal progenitor patent earlier with a narrower claim and then try to acquires a broader claim by a divisional application filed based on that patent application under circumstances where it wants to achieve both quick patent protection and eventually broader patent protection for an important invention.

Table 5 presents the summary results for the OLS and IV regression for the data of each of the four technology sectors. The descriptive statistics are given in Appendix II, Table II-1. According to Table 5, the coefficients of *ln_inverse_claim_length* and *ln_N_claims* are positive and highly significant in all technology sectors and the estimated coefficients are broadly similar across sectors. In addition, the estimated coefficient for q is also negative and statistically significant in all technology sectors, although it is marginally significant in *Others*. These results support both Hypotheses 1 and 2. The weaker statistical significance of the coefficient of q is likely to reflect the fact that the instrumental variables are based on companies and not on patents, so that the effective sample size is much smaller for estimation of the coefficient of q. The forward citation and the number of inventors have positive and significant coefficients, as estimated by IV estimations. The coefficients obtained by IV estimations for *US_Grant_dummy* are also positive and significant in E&E and Mech. These results support Hypothesis 3.

Table 5 here.

5.2. Extent of narrowing

Table 6 presents descriptive statistics and Table 7 a summary of the results of OLS and IV regressions of the extent of narrowing against *q*, the number of claims, the variables to measure the importance of the inventions, and the other controlling variables with fixed effects of applicant dummies, based on the pooled data. The coefficient of the *ln_N_claims* is positive and highly statistically significant at the 1% level for OLS and the two IV second stage estimations. The estimated coefficients for the three estimates are almost same. Thus, it is evident that an applicant who seeks a greater number of claims is more likely to have his/her first claim undergo a greater extent of narrowing as a result of patent examination, supporting Hypothesis 1. The elasticity with respect to the number of claims is large (around 0.18): a one standard deviation reduction in the number of claims (around 0.79 in logarithmic terms) results in around 14% points decrease in the scope of the rights through amendments made during the examination procedure.

Tables 6 and 7 here.

The coefficient for variable *q* is negative and statistically significant at the 1% level for all three estimates, indicating that for a higher-quality applicant's disclosure (where there is greater coverage of the examiner citations by the applicant citations), the extent of narrowing is smaller, supporting Hypothesis 2. An applicant who searches relevant prior art well before the patent application and designs the claim consistently with the prior art is more likely to acquire a scope of patent protection closer to what was requested at filing. Similar to the estimation results for the incidence of narrowing, IV estimation gives a more than seven times larger coefficient for the effect of disclosure quality, *q*, suggesting the importance of controlling for the measurement error of disclosure quality.

The coefficient for the *ln_N_Forward_Citation* has a positive and statistically significant value of around 0.02, supporting Hypothesis 3. However, the coefficient for the *US_Grant_dummy* is negative and statistically significant (around -1.1 to -0.8), unlike that for the incidence of narrowing; and there are no significant differences between the OLS and the two IV results. The different signs of the coefficient for *US_Grant_dummy* for the narrowing event and for the extent of narrowing imply that an application that is also filed in the US undergoes less restriction of the claims (i.e., the applicant needs to add fewer restrictions to the claim), even though a narrowing event itself is more likely to happen, after controlling for the number of claims, which is a proxy variable for the patent scope in the estimation model of the extent of narrowing. Our interpretation of this result is as follows. As the cost of filing a patent application abroad is very high, such an application would not only be more important but also the applicant would tend to search the prior art more thoroughly to avoid incurring such cost for a failed foreign application, and a thorough search would make the applicant to prepare the claim so that it more closely describes the contribution of the invention. Thus, *US_Grant_dummy* is correlated not only with higher v but also with lower θ. The former effect (the effect of higher v) dominates for the incidence of narrowing but the latter effect (the effect of lower θ) dominates for the extent of narrowing, because the latter estimation is based on a more selective sample of patent applications that are relatively more important[13], and the marginal effect of v decreases for more important patents unlike that of θ, according to equation (7). The coefficient for the number of inventors is negative (around -0.02) and highly significant in the OLS estimates, unlike the coefficient for the incidence of narrowing, but it is smaller and insignificant in the two IV estimates. Our interpretation of this is similar to that for the US application dummy. The number of inventors is correlated not only with higher v but also with lower θ because a larger number of inventors enables

[13] The mean values of all patent value indicator variables (the number of inventors, the forward citations and the incidence of the US grant) are larger for the sample for the estimation of the extent of narrowing than that for the estimation of the incidence of the narrowing (see Table 3 and 6).

discoveries of more relevant prior art. Such an effect can lead to a reduced effect of the number of inventors on the extent of narrowing.

Table 8 presents the summary results of IV regression for each of the four technology sectors. The descriptive statistics are given in Appendix II, Table II-2. According to Table 8, the coefficient of the *ln_N_claims* is positive and highly statistically significant in all technology sectors and the estimated coefficients are broadly similar across sectors. The estimated coefficient for q is negative and statistically significant in all technology sectors except for C & C, where the estimated coefficient for q is negative but statistically insignificant. These results support both Hypothesis 1 and 2. Again, the weaker significance of the coefficient of q is likely to reflect the fact that the instrumental variables are based on companies and not on patents, so that the effective sample size is much smaller. In C & C, the average of q is 0.034; this is much smaller than for the other fields (E & E, 0.046; Mech., 0.064; and Others, 0.051, respectively). The estimated coefficient for q is also not statistically significant and the absolute value is smallest for C & C. These results imply that, for applicants in the C & C field, it is difficult both to find the relevant prior art and to predict the claim that satisfies such a requirement as an inventive step over the closest prior art.

Table 8 here.

5.3. Robustness check

As a robustness check, we performed estimations using the number of prior patents cited by the applicant as a measure of disclosure in place of the prior art disclosure quality variable q, which depends partly on the examiners. As the applicant increases the number of prior art citations, we would expect the disclosure quality also to improve. The results for the incidence of narrowing are shown in Appendix III, Table III-1 (descriptive statistics), and Table III-2 (estimation results). The qualitative results are the same as those presented in Table 4. According to the estimation results, a patent application with greater disclosure is less likely to undergo restriction of the scope of the first claim, and that broader claims (according to the inverse of the initial claim length and the number of claims) are more likely to experience restrictions.

The results for the extent of narrowing are shown in Appendix III, Table III-3 (descriptive statistics), and Table III-4 (estimation results). The qualitative results are the same as those presented in Table 7. According to the estimation results, applicants that disclose more prior art before patent applications are more likely to acquire a scope of patent protection closer to that requested at filing. Applicants seeking broader claims (i.e., more claims) are more likely to experience a larger extent of narrowing of their first claims.

6. Conclusions

Delineating the scope of a patent so that it is consistent with the contribution of the disclosed invention to the state of the art is one of the most crucial requirements for a patent system to promote innovation effectively. It is a key element of high patent quality. Given the incentive of an applicant to claim a patent right with the broadest scope possible, an important task of the Patent Office is to narrow overly broad claims so that they become consistent with the contribution of the filed invention to the state of the art. This study analyzed empirically to what extent the Patent Office delivers this function through patent examination, focusing on product patents in four major technology areas. To the best of our knowledge, the economic literature has not reported a full-scale econometric analysis of this issue. In this paper, we developed a simple model of patent examination and investigated how the breadth of the initial claim at the time of filing, the quality of the applicant's disclosure of relevant prior art, and the importance of the patent application drive the incidence and the extent of narrowing of the initial claim, in order to assess the key predictions of our model. We used the inverse of the description length of the first claim as well as the number of claims as proxy variables for the scope of a patent right and focused on product inventions in four major technology areas where cascaded-type claims are dominant.

We found that the first claims were narrowed as a result of the patent examination for two-thirds of the granted patents in all four major technology areas. Such an adjustment was significantly less likely when the applicant disclosed prior art more relevant to the filed invention, controlling for the endogeneity, but was more likely when the applicant chose a broad initial scope at filing (i.e., the application had the first claim with a short description or a large number of claims). Similarly, if an amendment is made, the extent of narrowing decreased with a quality of prior art disclosure and increased with an initial claim scope. We also found that patent applications that were more important in terms of the number of inventors, the forward citations, and the existence of corresponding US patents, underwent narrowing more frequently. Thus, patent examination makes a major contribution to ensuring that claims are consistent with the contribution of the filed inventions to the state of the art, especially when the economic impact is likely to be large.

These main results hold across the board for the four major technology fields where cascade-type claiming is considered to be dominant, i.e., C & C, E & E, Mech., and Others. Out of the four, C & C has the lowest average quality of disclosure of prior art by the applicant. The quality of prior art disclosure has no statistically significant effect on the extent of narrowing. These facts imply that it is difficult both to find the relevant prior art and to predict the claim that satisfies such a requirement as an inventive step over the closest prior art.

Although more important patent applications (in terms of the number of inventors, the forward citations, and the existence of corresponding US patents) undergo narrowing process more frequently, the effect on the extent of narrowing varies depending on the indicators of the importance of the invention. For an example, the existence of a US corresponding patent is found to have a negative effect on the extent of narrowing. This contrasting result seems to indicate that an applicant is likely to prepare a better quality patent application with a better prior art disclosure and with more appropriately elaborated claim when the patent application cost is high. This issue deserves further research.

The results of this study show that the patent examiner plays a very significant role in delineating the scope of a patent, which is a cornerstone of the patent system in promoting innovation. This reaffirms the importance of patent examination for ensuring high patent quality. It will be important to develop infrastructure for supporting the Patent Office to effectively search and access prior art, and to engage in international cooperation in search and examination; the effective use of third-party contributions such as information submissions, and post-grant oppositions will also be important.

Furthermore, our empirical result that more important patent applications undergo narrowing process more frequently are consistent with the predictions of our model of patent examination in which there is a time constraint for patent examiners so that they allocate more resources to important patent applications in order to avoid a large unjustified economic loss such as an excess deadweight loss due to an excessive claim. There is a question of how the examiners can be motivated to pursue such efficient examinations, i.e., to devote more resources to important patent applications. As there has been no major financial incentives for this objective and furthermore, there is pressure on examiners to increase the number of patent examinations processed and the speed of examinations, an intrinsic motivation and/or a sense of responsibility seems to play a major role in the efficient allocation of each examiner's time. Future research on an incentive system to sustain and enhance such non-financial incentives as intrinsic motivation would seem to be important.

The results also suggest that policies that support improvement of the quality of prior art search by applicants will make a difference to patent examination efficiency. Improving applicants' disclosure of prior art, being accompanied with claim drafting consistent with the contribution of the filed inventions to the state of the art, will cause a reduction in the workload of the Patent Office (for example, by requiring less frequent amendments) and will increase the average amount of time that can be allocated to the examination of each patent application; thus it will improve the quality of patents as a whole. It is therefore important for the Patent Office to make prior art information easily accessible by the applicants through improving the infrastructure of its patent information search system. In C & C, it seems that applicants face difficulty in both finding the closest prior art and drafting the broadest claims that satisfy patentability requirements. High priority should be given to improving the prior art search infrastructures available to applicants in this field.

Acknowledgment

This study is conducted as a part of the project “Study on Industrial Innovation Capability and Innovation Infrastructure” undertaken at Research Institute of Economy, Trade and Industry (RIETI). The earlier stage of this work was supported by JSPS KAKENHI, Grant Number 26885031 and 26285055. We would like to thank JSPS for granting this aid. We would like to thank Hiroshi Ohashi, Atsushi Ohyama, Tetsuo Wada, Gaétan de Rassenfosse, Adam Jaffe, and Stuart Graham, the participants at RIETI seminars, the IP Statistics for Decision Makers Conference 2015, the Asia–Pacific Innovation Conference 2015 and 2016, Innovation Economics Workshop 2016 the second, Japanese Economic Association Spring meeting 2016, the 2016 Asian Meeting of the Econometric Society, European Association for Research in Industrial Economics 2016, and European Policy for Intellectual Property 2016 for their helpful comments on earlier versions of this study. We also acknowledge research assistant Junko Morikawa for her help in the sample check investigation on category data of the first claims.

References

Abiko, G., & Tanaka, Y. (2010), The Study of Variance of Decision Making to Claim Influence on Usefulness, Journal of Intellectual Property Association of Japan, 7(1), 54-67.

Bessen, J., & Meurer, M. J. (2008). Patent failure: How judges, bureaucrats, and lawyers put innovators at risk. Princeton University Press.

Cotropia, C. A., Lemley, M. A., & Sampat, B. (2013). Do applicant patent citations matter?. Research Policy, 42(4), 844-854.

Farrell, Joseph and Carl Shapiro. (2008). "How Strong Are Weak Patents?" American Economic Review, 98(4), 1347-1369.

Federal Trade Commission (FTC). (2003) To Promote Innovation: The Proper Balance of Competition and Patent Law and Policy

Goto A., & Motohashi, K. (2007), Construction of a Japanese Patent Database and a first look at Japanese patenting activities, Research Policy, 36(9), 1431-1442.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001), The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools, National Bureau of Economic Research, No. w8498.

Harhoff, D. (2016). Patent Quality and Examination in Europe. The American Economic Review: Papers & Proceedings 2016, 106(5), 193-197.

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

Lampe, R. (2012). Strategic citation. Review of Economics and Statistics, 94(1), 320-333.

Lemley, M. A., & Shapiro, C. (2005). Probabilistic patents. The Journal of Economic Perspectives, 19(2), 75-98.

Malackowski, J., & Barney, J. (2008), What is Patent Quality? A Merchant Banc's Perspective, Les Nouvelles, June, pp. 123-134. Retrieved from: http://www.lesi.org/les-nouvelles/les-nouvelles-online/2008/june-2008/2011/08/05/what-is-patent-quality-a-merchant-banc-s-perspective

National Academies (2004) A Patent System for the 21st Century

NISTEP, 2015, "Company name dictionary."

URL: http://www.nistep.go.jp/research/scisip/data-and-information-infrastructure (accessed April 17, 2015).

NISTEP, 2015, "Connection table to other databases."

URL: http://www.nistep.go.jp/research/scisip/data-and-information-infrastructure (accessed April 17, 2015).

Okada, Y., Naito Y. & Nagaoka S. (2016), Claim Length as a Value Predictor of a Patent, IIR Working Paper 16-04

Osenga, K. (2011), Shape of Things to Come: What We Can Learn from Patent Claim Length, Santa Clara Computer & High Tech, LJ 28, pp. 617-656.

Sampat, B. N. (2010). When do applicants search for prior art? Journal of Law and Economics, 53(2), 399-416.

Schuett, F. (2013). Patent quality and incentives at the patent office. The RAND Journal of Economics, 44(2), 313-336.

US Government Accountability Office (GAO) (2013) INTELLECTUAL PROPERTY: Assessing Factors That Affect Patent Infringement Litigation Could Help Improve Patent Quality, GAO-13-465

Webster E., Jensen P. H., & Palangkaraya A. (2014), Patent Examination Outcomes and the National Treatment Principle, RAND Journal of Economics, 45(2), 449-469.

Figure 1 Marginal benefit and cost for a patent examiner of narrowing the scope

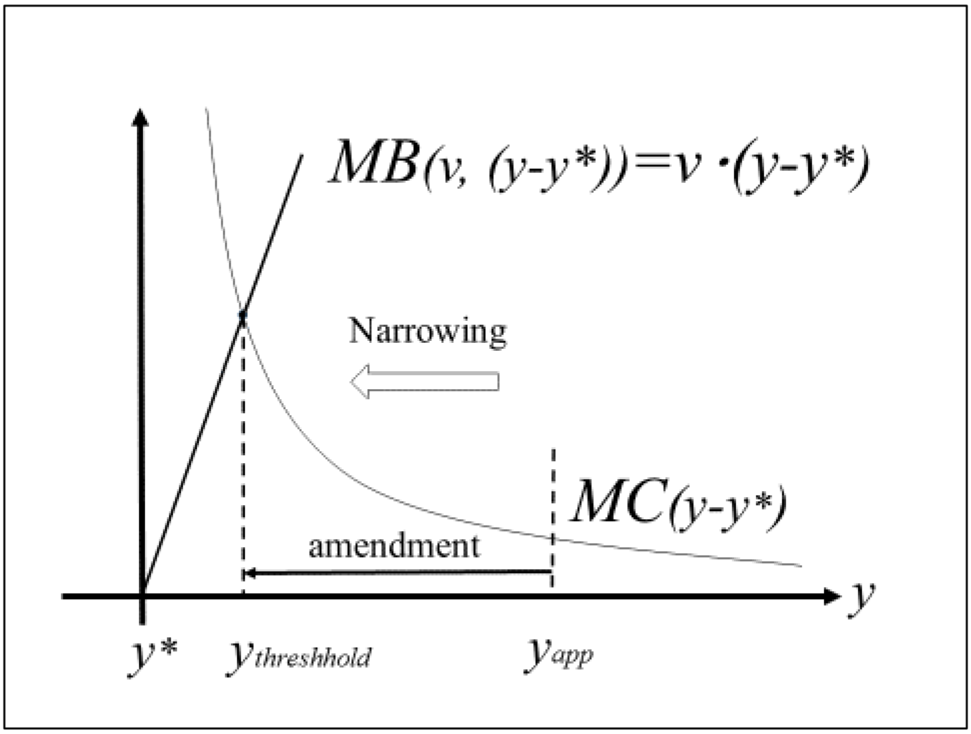


Figure 2 Distribution of ln(claim length) of granted patents for product inventions

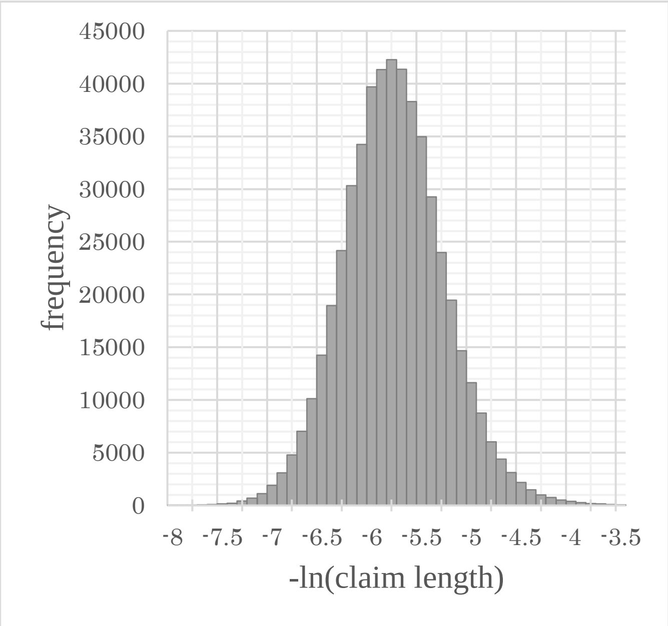
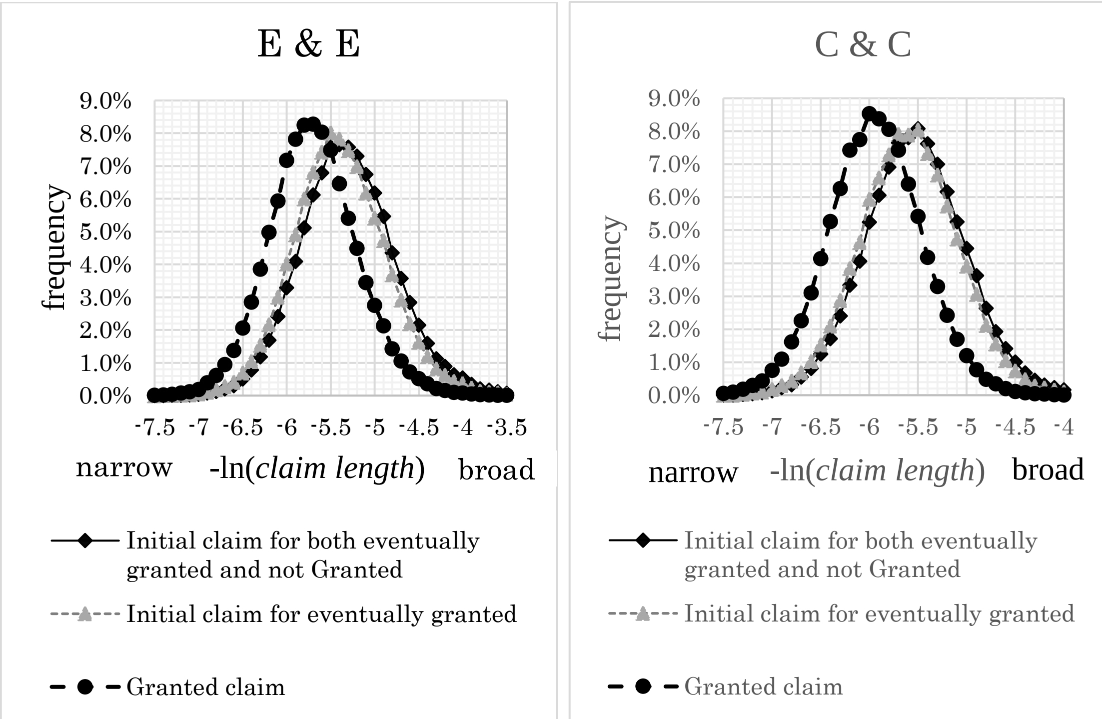
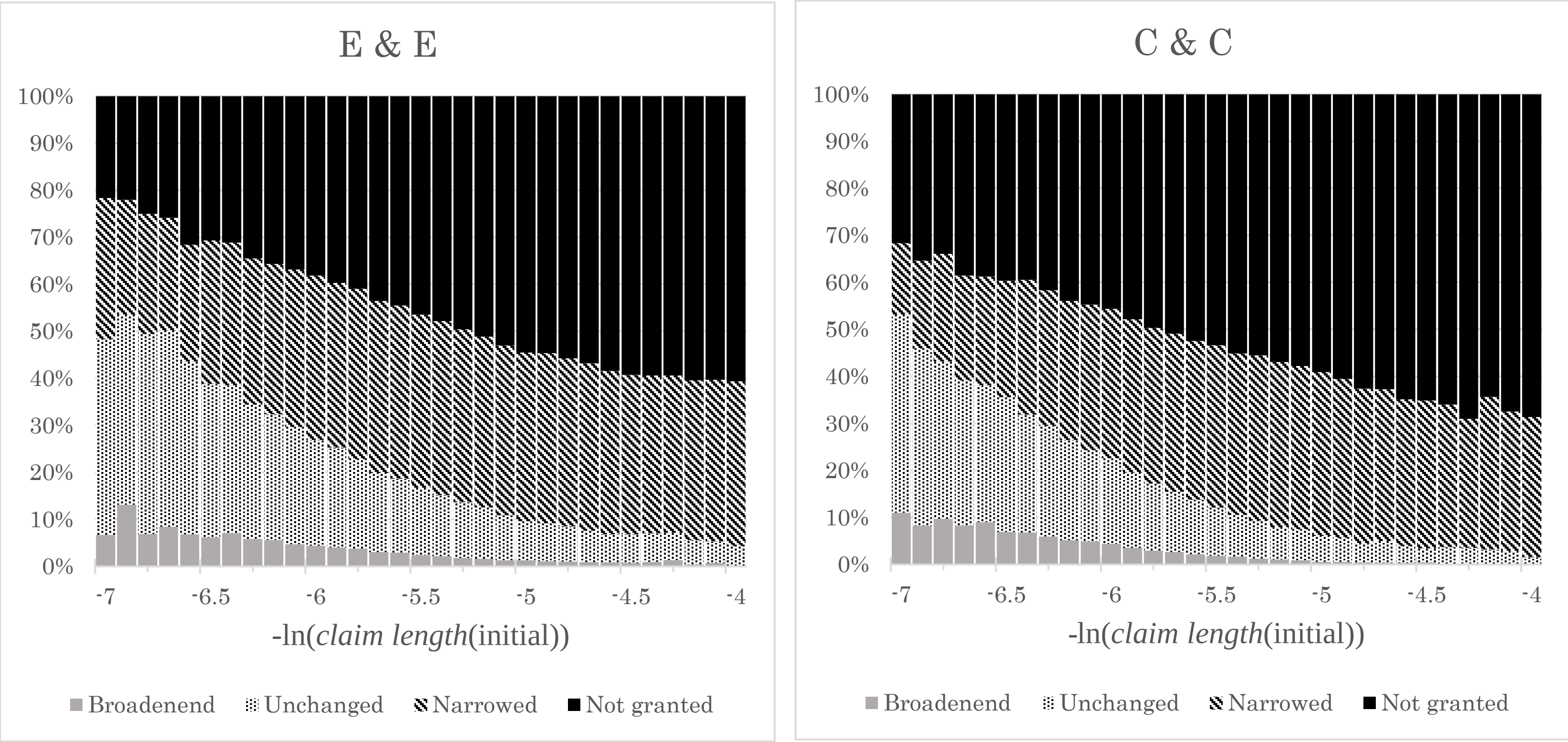


Figure 3 Comparison of the distribution of the natural logarithm of the inverse of the claim length before and after examination



Note: Left: Electrical & Electronic (E & E) and right: Computers & Communications (C & C).

Figure 4 Distribution of -ln(*claim length*) and the outcome of the examination



Note: Left: Electrical & Electronic (E & E) and right: Computers & Communications (C & C). “Not granted” includes rejected, withdrawn, abandoned, and dismissed after the first action of the patent examiner. “Broadened” may cover cases where the scope was reduced by eliminating choices or alternatives written in an initial parallel-type claim.

Figure 5 The quality of an applicant's disclosure of prior art documents

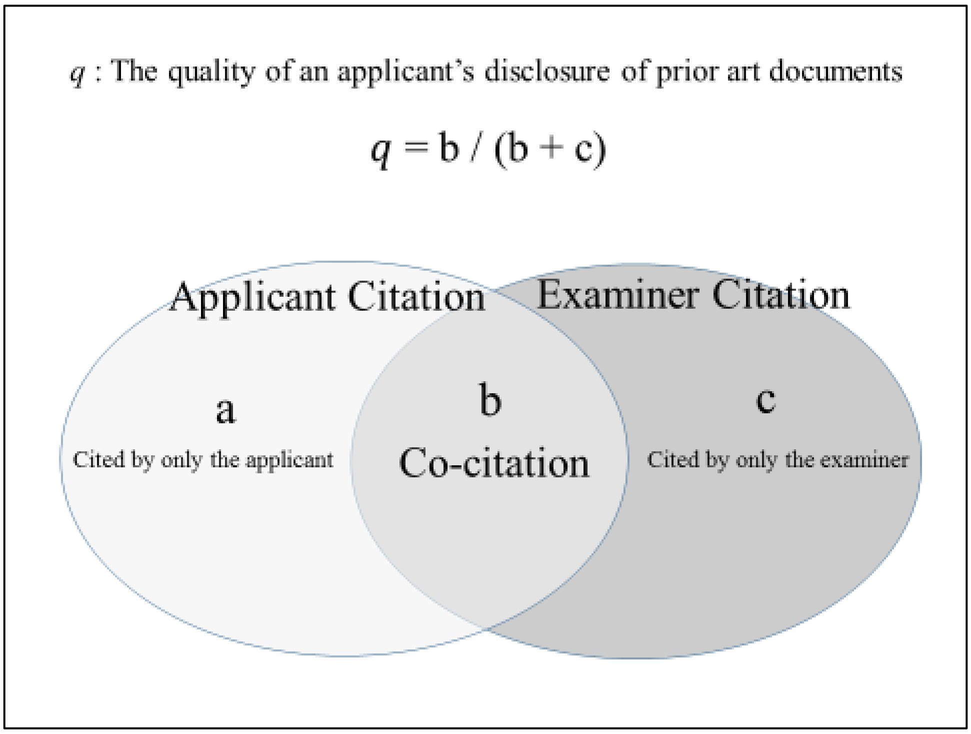


Table 1 Claim length (the number of characters in the first claim) of granted patents

Technology field	Category of invention	N	mean	sd	p1	p5	med	p95	p99
C & C	product	125255	429.1	224.4	119	172	382	842	1195
	process	21899	424.2	235.1	120	169	370	859	1245
E & E	product	152288	336.7	174.6	84	129	303	659	920
	process	39862	320.3	169.8	88	126	284	632	901
Mech.	product	145197	372.5	186.4	96	145	337	718	990
	process	24459	317.3	171.1	91	127	281	633	910
Others	product	95263	316.5	170.8	63	110	284	631	888
	process	16935	260.8	148.6	59	89	230	534	774

Note: Left: C & C, E & E, and Mech. stands for Computers & Communications, Electrical & Electronic, and Mechanical, respectively.

Table 2A Composition of the outcome of patent examinations

Technology field	Granted "broadened"	Granted unchanged	Granted narrowed	Not granted	Total
C & C	6,355	31,227	84,680	118,165	240,427
E & E	6,663	39,427	103,438	119,734	269,262
Mech.	4,466	38,423	99,596	99,398	241,883
Others	2,839	29,831	60,886	63,312	156,868
Total	20,323	138,908	348,600	400,609	908,440

Note: C & C, E & E, and Mech. stands for Computers & Communications, Electrical & Electronic, and Mechanical, respectively. "Not granted" includes rejected, withdrawn, and abandoned after first action of patent examiner. "Broadened" may cover cases where the scope was reduced by eliminating choices or alternatives written in an initial parallel-type claim.

Table 2B Composition ratio of the outcome of patent examination

Technology field	Granted	Granted "broadened"	Granted unchanged	Granted narrowed	Not granted
C & C	50.9%	5.2%	25.5%	69.3%	49.1%
E & E	55.5%	4.5%	26.4%	69.2%	44.5%
Mech.	58.9%	3.1%	27.0%	69.9%	41.1%
Others	59.6%	3.0%	31.9%	65.1%	40.4%
Total	55.9%	4.0%	27.4%	68.6%	44.1%

Note: C & C, E & E, and Mech. stands for Computers & Communications, Electrical & Electronic, and Mechanical, respectively. "Not granted" includes rejected, withdrawn, abandoned, and dismissed after the first action of the patent examiner. "Broadened" may cover cases where the scope was reduced by eliminating choices or alternatives written in an initial parallel-type claim.

Table 3 Descriptive statistics for the incidence of narrowing, based on pooled data

variable	mean	sd	p99	p1
incidence of narrowing	0.747	0.435	1	0
q = b / (b + c)	0.0538	0.157	1	0
av_q	0.0539	0.0351	0.159	0
ln_av_N_Backward_Citation	0.554	0.357	1.74	0.0328
ln_inverse_claim_length(initial)	-5.46	0.520	-4.14	-6.65
ln_N_claims(initial)	1.63	0.798	3.56	0
ln_N_inventor	0.590	0.605	1.95	0
ln_N_Forward_Citation	0.573	0.736	2.89	0
US_Grant_dummy	0.257	0.437	1	0
subAppln_dummy	0.0479	0.213	1	0

N = 326,859

Table 4 Summary results of ordinary least squares (OLS) and instrumental variables (IV) regression for the incidence and extent of narrowing, based on pooled data

	Dependent variable: *incidence of narrowing* (narrowed=1, unchanged=0)				
	(1)	(2)	(3)	(4)	(5)
	OLS	IV	IV	IV	IV
Independent variables		second stage	first stage	second stage	first stage
q = b/(b+c)	-.132***	-.695***		-.804***	
	(.00468)	(.135)		(.158)	
av_q			.176***		.230***
			(.0142)		(.0132)
ln_av_N_Backward_Citation			.0197***		
			(.00191)		
ln_inverse_claim_length(initial)	.189***	.179***	-.0174***	.177***	-.0174***
	(.00148)	(.00281)	(.000554)	(.00316)	(.000554)
ln_N_claims(initial)	.0472***	.0483***	.00184***	.0485***	.00186***
	(.00113)	(.00118)	(.000422)	(.00120)	(.000422)
ln_N_inventor	.00451***	.00671***	.00379***	.00714***	.00388***
	(.00133)	(.00146)	(.000498)	(.00151)	(.000498)
ln_N_Forward_Citation	.0182***	.0185***	.000470	.0186***	.000512
	(.00102)	(.00104)	(.000381)	(.00105)	(.000381)
US_Grant_dummy	.00538***	.00816***	.00475***	.00870***	.00486***
	(.00183)	(.00198)	(.000684)	(.00204)	(.000684)
subAppln_dummy	.104***	.101***	-.00551***	.100***	-.00551***
	(.00348)	(.00363)	(.00130)	(.00369)	(.00130)
effective filing year	yes	yes	yes	yes	yes
33_clasiffication	yes	yes	yes	yes	yes
*effective filing year * 33_clasiffication*	yes	yes	yes	yes	yes
applicant	FE	FE	FE	FE	FE
Observations	326,859	326,859	326,859	326,859	326,859
R-squared	.0805		.00794		.00761
Number of aplt_history_id1	260	260	260	260	260
Standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.10					

Table 5 Summary results of instrumental variables (IV) regression for the incidence of narrowing

	Dependent Variable: *incidence of narrowing* (narrowed=1, unchanged=0)											
	C & C			E & E			Mech.			Others		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	IV	OLS	IV	IV	OLS	IV	IV	OLS	IV	IV
Independent variables		second stage	first stage		second stage	first stage		second stage	first stage		second stage	first stage
q = b/(b+c)	-.174***	-1.28***		-.127***	-1.37***		-.117***	-1.79**		-.114***	-1.39*	
	(.0100)	(.459)		(.00879)	(.445)		(.00802)	(.831)		(.0118)	(.798)	
av_q			.229***			.200***			.0891***			.106***
			(.0321)			(.0288)			(.0256)			(.0304)
ln_inverse_claim_length(initial)	.205***	.192***	-.0121***	.166***	.146***	-.0160***	.191***	.154***	-.0221***	.208***	.178***	-.0231***
	(.00273)	(.00626)	(.000886)	(.00260)	(.00767)	(.000921)	(.00297)	(.0187)	(.00126)	(.00425)	(.0191)	(.00172)
ln_N_claims(initial)	.0433***	.0451***	.00158**	.0434***	.0466***	.00260***	.0506***	.0540***	.00204**	.0639***	.0657***	.00134
	(.00202)	(.00228)	(.000657)	(.00197)	(.00243)	(.000699)	(.00228)	(.00328)	(.000968)	(.00343)	(.00403)	(.00139)
ln_N_inventor	.00247	.00618**	.00332***	.00287	.00878***	.00469***	.00212	.00952*	.00443***	.0148***	.0175***	.00208
	(.00257)	(.00314)	(.000835)	(.00232)	(.00330)	(.000824)	(.00260)	(.00486)	(.00110)	(.00369)	(.00449)	(.00150)
ln_N_Forward_Citation	.0215***	.0224***	.000743	.0164***	.0181***	.00135**	.0140***	.0125***	-.000903	.0286***	.0285***	-.0000919
	(.00198)	(.00213)	(.000642)	(.00167)	(.00193)	(.000594)	(.00199)	(.00256)	(.000846)	(.00318)	(.00358)	(.00129)
US_Grant_dummy	.000837	.00382	.00272***	.0132***	.0210***	.00626***	.00701*	.0142**	.00419***	-.00916	.00405	.0104***
	(.00313)	(.00355)	(.00102)	(.00303)	(.00433)	(.00108)	(.00366)	(.00572)	(.00156)	(.00881)	(.0129)	(.00358)
subAppln_dummy	.101***	.0955***	-.00496***	.107***	.0963***	-.00852***	.100***	.0954***	-.00292	.106***	.105***	-.000784
	(.00571)	(.00649)	(.00186)	(.00596)	(.00751)	(.00212)	(.00788)	(.00998)	(.00335)	(.0111)	(.0125)	(.00451)
effective filing year	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
33_clasiffication	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
*effective filing year * 33_clasiffication*	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
applicant	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Observations	94,287	94,287	94,287	102,639	102,639	102,639	86,166	86,166	86,166	43,767	43,767	43,767
R-squared	.0855		.00351	.0712		.00509	.0786		.00603	.0954		.00930
Number of applicant	75	75	75	97	97	97	121	121	121	86	86	86
Standard errors in parentheses												
*** p<0.01, ** p<0.05, * p<0.10												

Table 6 Descriptive statistics for the extent of narrowing, based on pooled data

variable	mean	sd	p99	p1
extent of narrowing	-6.75	1.20	-4.51	-10.57
q = b / (b + c)	0.0480	0.145	0.75	0
av_q	0.0532	0.0345	0.156	0
ln_av_N_Backward_Citation	0.560	0.360	1.74	0.0333
ln_N_claims(initial)	1.70	0.790	3.61	0
ln_N_inventor	0.601	0.610	2.08	0
ln_N_Forward_Citation	0.603	0.752	2.94	0
US_Grant_dummy	0.269	0.443	1	0
subAppln_dummy	0.0591	0.236	1	0

N = 244,029

Table 7 Summary results of ordinary least squares (OLS) and instrumental variable (IV) regression for the extent of narrowing, based on pooled data

	Dependent Variable: *extent of narrowing* ln(inverse claim length(initial) - inverse claim_length(grant))				
	(1)	(2)	(3)	(4)	(5)
	OLS	IV	IV	IV	IV
VARIABLES		second stage	first stage	second stage	first stage
q = b/(b+c)	-.466***	-4.01***		-4.56***	
	(.0165)	(.485)		(.567)	
av_q			.183***		.234***
			(.0157)		(.0146)
ln_av_N_Backward_Citation			.0177***		
			(.00202)		
ln_N_claims(initial)	.178***	.181***	.00104**	.182***	.00104**
	(.0037)	(.00407)	(.000455)	(.00418)	(.000455)
ln_N_inventor	-.0125***	-.00152	.00297***	.000194	.00305***
	(.0043)	(.00493)	(.000529)	(.00513)	(.000529)
ln_N_Forward_Citation	.0201***	.0225***	.000608	.0229***	.00065
	(.00325)	(.00356)	(.000399)	(.00366)	(.000399)
subAppln_flg	.122***	.107***	-.00422***	.105***	-.00421***
	(.0103)	(.0114)	(.00126)	(.0118)	(.00126)
US_Grant_flg	-.106***	-.0842***	.00592***	-.0808***	.00602***
	(.00588)	(.00706)	(.000723)	(.00744)	(.000723)
effective filing year	yes	yes	yes	yes	yes
33_clasiffication	yes	yes	yes	yes	yes
*effective filing year * 33_clasiffication*	yes	yes	yes	yes	yes
applicant	FE	FE	FE	FE	FE
Observations	244,029	244,029	244,029	244,029	244,029
R-squared	.0247		.00531		.00500
Number of aplt_history_id1	260	260	260	260	260
Standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.10					

Table 8 Summary results of instrumental variables (IV) regression for the extent of narrowing

	Dependent Variable: *extent of narrowing* ln(inverse claim length(initial) - inverse claim_length(grant))											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	C & C			E & E			Mech.			Others		
	OLS	IV	IV	OLS	IV	IV	OLS	IV	IV	OLS	IV	IV
VARIABLES		second stage	first stage		second stage	first stage		second stage	first stage		second stage	first stage
q = b/(b+c)	-.429***	-2.12		-.524***	-3.99***		-.439***	-6.73**		-.421***	-8.74**	
	(.0371)	(2.27)		(.0311)	(1.19)		(.0273)	(2.87)		(.0397)	(3.42)	
av_q			.152***			.250***			.093***			.109***
			(.0346)			(.0319)			(.0286)			(.0343)
ln_N_claims(initial)	.145***	.147***	.000902	.179***	.184***	.00181**	.204***	.209***	.000748	.203***	.211***	.000881
	(.00676)	(.00717)	(.000686)	(.00655)	(.00734)	(.000758)	(.00733)	(.0101)	(.00106)	(.0108)	(.0170)	(.00153)
ln_N_inventor	-.0178**	-.0137	.00238***	.0137*	.0264***	.00359***	-.0240***	-.00505	.00303**	-.0281**	-.00339	.00292*
	(.00844)	(.0102)	(.000855)	(.00760)	(.00926)	(.000880)	(.00828)	(.0142)	(.00120)	(.0113)	(.0203)	(.00162)
ln_N_Forward_Citation	.0260***	.0281***	.00119*	.0195***	.0248***	.00149**	.0177***	.0119	-.000922	.0199**	.0197	-.000036
	(.00642)	(.00708)	(.000651)	(.00541)	(.0061)	(.000626)	(.00628)	(.00889)	(.000907)	(.00957)	(.0148)	(.00136)
US_Grant_dummy	-.0749***	-.0696***	.00311***	-.112***	-.0859***	.00751***	-.129***	-.0892***	.00625***	-.112***	-.0214	.0110***
	(.0103)	(.0126)	(.00105)	(.00989)	(.0139)	(.00114)	(.0115)	(.0240)	(.00166)	(.0269)	(.0559)	(.00383)
subAppln_dummy	.128***	.123***	-.00340*	.0955***	.0730***	-.00672***	.162***	.147***	-.00240	.0951***	.0969**	.000180
	(.0174)	(.0192)	(.00176)	(.0179)	(.0208)	(.00207)	(.0229)	(.0317)	(.00331)	(.0307)	(.0475)	(.00437)
effective filing year	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
33_clasiffication	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
*effective filing year * 33_clasiffication*	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
applicant	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Observations	70,790	70,790	70,790	77,255	77,255	77,255	64,527	64,527	64,527	31,457	31,457	31,457
R-squared	.0183		.00141	.0178		.0026	.0279		.00195	.0348		.00686
Number of aplt_history_id1	75	75	75	97	97	97	121	121	121	86	86	86
Standard errors in parentheses												
*** p<0.01, ** p<0.05, * p<0.10												

Appendix I Classification of technology sectors

Table I-1

Technology sector	Detailed 33 clasification	IPC
Computers & Communications	Clock, Controling, Computer	G04−G08
Computers & Communications	Display, Information Storage, Instruments	G09−G12
Computers & Communications	Electronics circuit, Communication tech.	H03−H04
Electrical & Electronic	Measurement, Optics, Photography	G01−G03
Electrical & Electronic	Electronics components, Semiconductor	H01−H02,H05
Mechanical	Machine tools, Metal working	B21−B23
Mechanical	Casting, Grinding, Layered product	B24−B32(excludes B31)
Mechanical	Printing	B41−B44
Mechanical	Transporting	B60−B64
Mechanical	Packing, Lifting	B65−B68
Mechanical	Engine, Pump	F01−F04,F15
Mechanical	Engineering elements	F16−F17
Chemical	Nonorganic chemistry, Fertilizer	C01−C05
Chemical	Organic chemistry, Pesticides	C07,A01N
Chemical	Organic molecule compounds	C08
Chemical	Dyes, Petroleum	C09−C11
Chemical	Metallurgy, Coating metals	C21−C30
Drugs & Medical	Health and Amusument	A61−A63(excludes A61K)
Drugs & Medical	Drugs	A61K
Drugs & Medical	Biothecnology, Beer, Fermentation	C12−C14
Drugs & Medical	Genetic Engineering	C12N15/
Others	Agriculture	A01(excludes A01N)
Others	Food stuffs	A21−A24
Others	Personal and Domestic Articles	A41−A47
Others	Separating, Mixing	B01−B09
Others	Others	B81,B82
Others	Textile	D01−D07
Others	Paper	D21,B31
Others	Construction	E01−E06
Others	Mining, Drilling	E21
Others	Lighting, Steam generation, Heating	F21−F28
Others	Wepons, Blasting	F41−F42,C06
Others	Nuclear physics	G21

Appendix II Descriptive statistics for data used in instrumental variables (IV) estimation, by technology sector

Table II-1 Descriptive statistics for the IV estimation reported in Table 5

C & C	N=94,287			
variable	mean	sd	p99	p1
incidence of narrowing	0.751	0.433	1	0
q = b/(b+c)	0.0394	0.134	0.667	0
av q	0.0390	0.0256	0.118	0
ln av N Backward Citation	0.457	0.255	1.15	0.0172
ln_inverse_claim_length(initial)	-5.57	0.500	-4.37	-6.75
ln N claims(initial)	1.85	0.789	3.76	0
ln N inventor	0.464	0.583	1.95	0
ln N Forward Citation	0.539	0.695	2.64	0
US Grant dummy	0.316	0.465	1	0
subAppln dummy	0.0615	0.240	1	0
E & E	N=102,639			
variable	mean	sd	p99	p1
incidence of narrowing	0.753	0.431	1	0
q = b/(b+c)	0.0505	0.148	0.800	0
av q	0.0505	0.0269	0.134	0
ln av N Backward Citation	0.667	0.441	2.70	0.0522
ln inverse claim length(initial)	-5.38	0.516	-4.06	-6.55
ln N claims(initial)	1.71	0.793	3.58	0
ln N inventor	0.620	0.612	2.08	0
ln N Forward Citation	0.634	0.797	3.14	0
US Grant dummy	0.308	0.462	1	0
subAppln dummy	0.0521	0.222	1	0
Mech.	N=86,166			
variable	mean	sd	p99	p1
incidence of narrowing	0.749	0.434	1	0
q = b/(b+c)	0.0710	0.177	1	0
av q	0.0718	0.0384	0.174	0
ln av N Backward Citation	0.591	0.310	1.61	0.0455
ln inverse claim length(initial)	-5.50	0.508	-4.22	-6.64
ln N claims(initial)	1.40	0.761	3.22	0
ln_N_inventor	0.603	0.590	1.95	0
ln N Forward Citation	0.592	0.734	2.83	0
US Grant dummy	0.227	0.419	1	0
subAppln dummy	0.0340	0.181	1	0
Others	N=43,767			
variable	mean	sd	p99	p1
incidence of narrowing	0.719	0.450	1	0
q = b/(b+c)	0.0587	0.174	1	0
av q	0.0587	0.0449	0.199	0
ln_av_N_Backward_Citation	0.426	0.309	1.39	0.0174
ln inverse claim length(initial)	-5.32	0.532	-3.95	-6.51
ln N claims(initial)	1.39	0.718	3.04	0
ln N inventor	0.761	0.613	2.08	0
ln N Forward Citation	0.465	0.657	2.56	0
US Grant dummy	0.067	0.250	1	0
subAppln dummy	0.0356	0.185	1	0

Table II-2 Descriptive statistics for the IV estimation reported in Table 8

C & C	N=70,790			
variables	mean	sd	p99	p1
extent of narrowing	-6.79	1.20	-4.66	-10.72
q = b/(b+c)	0.0337	0.121	0.500	0
av_q	0.038	0.025	0.12	0
ln_N_claims(initial)	1.9197	0.7833	3.829	0
ln_N_inventor	0.478	0.590	1.95	0
ln_N_Forward_Citation	0.571	0.712	2.71	0
US_Grant_dummy	0.328	0.470	1	0
subAppln_dummy	0.08	0.264	1	0
E & E	N=77,255			
variables	mean	sd	p99	p1
extent of narrowing	-6.72	1.20	-4.45	-10.52
q = b/(b+c)	0.0461	0.138	0.667	0
av_q	0.0504	0.0267	0.134	0
ln_N_claims(initial)	1.78	0.779	3.64	0
ln_N_inventor	0.635	0.617	2.08	0
ln_N_Forward_Citation	0.666	0.814	3.22	0
US_Grant_dummy	0.321	0.467	1	0
subAppln_dummy	0.0642	0.245	1	0
Mech	N=64,527			
variables	mean	sd	p99	p1
extent of narrowing	-6.83	1.19	-4.59	-10.58
q = b/(b+c)	0.0643	0.166	1.000	0
av_q	0.0709	0.0378	0.169	0
ln_N_claims(initial)	1.47	0.755	3.30	0
ln_N_inventor	0.610	0.593	1.95	0
ln_N_Forward_Citation	0.615	0.745	2.89	0
US_Grant_dummy	0.238	0.426	1	0
subAppln_dummy	0.0417	0.200	1	0
Others	N=31,457			
variables	mean	sd	p99	p1
extent of narrowing	-6.53	1.17	-4.33	-10.20
q = b/(b+c)	0.0518	0.160	1.00	0
av_q	0.0577	0.0440	0.195	0
ln_av_N_Backward_Citation	0.426	0.304	1.37	0.0211
ln_N_inventor	0.778	0.618	2.08	0
ln_N_Forward_Citation	0.493	0.675	2.64	0
US_Grant_dummy	0.069	0.253	1	0
subAppln_dummy	0.0452	0.208	1	0

Appendix III Estimations based on the number of disclosures by the applicants

Table III-1 Descriptive statistics for the estimations reported in Table III-2

C & C	N=115,889			
variable	mean	sd	p99	p1
incidence of narrowing	0.731	0.444	1	0
ln_N_Backward_Citation	0.330	0.498	1.79	0
ln_inverse_claim_length(initial)	-5.58	0.506	-4.34	-6.76
ln_N_claims(initial)	1.79	0.800	3.71	0
ln_N_inventor	0.464	0.577	1.95	0
ln_N_Forward_Citation	0.529	0.689	2.64	0
US_Grant_dummy	0.293	0.455	1	0
subAppln_dummy	0.0573	0.232	1	0
E & E	N=142,831			
variable	mean	sd	p99	p1
incidence of narrowing	0.724	0.447	1	0
ln_N_Backward_Citation	0.448	0.629	2.56	0
ln_inverse_claim_length(initial)	-5.38	0.524	-4.03	-6.57
ln_N_claims(initial)	1.62	0.788	3.50	0
ln_N_inventor	0.606	0.599	1.95	0
ln_N_Forward_Citation	0.618	0.786	3.14	0
US_Grant_dummy	0.273	0.445	1	0
subAppln_dummy	0.0462	0.210	1	0
Mech	N=138,010			
variable	mean	sd	p99	p1
incidence of narrowing	0.722	0.448	1	0
ln_N_Backward_Citation	0.419	0.541	2.08	0
ln_inverse_claim_length(initial)	-5.48	0.521	-4.16	-6.65
ln_N_claims(initial)	1.35	0.747	3.14	0
ln_N_inventor	0.570	0.573	1.95	0
ln_N_Forward_Citation	0.552	0.712	2.77	0
US_Grant_dummy	0.191	0.393	1	0
subAppln_dummy	0.0304	0.172	1	0
Others	N=90,708			
variable	mean	sd	p99	p1
incidence of narrowing	0.671	0.470	1	0
ln_N_Backward_Citation	0.340	0.523	2.08	0
ln_inverse_claim_length(initial)	-5.30	0.571	-3.74	-6.53
ln_N_claims(initial)	1.32	0.714	3.00	0
ln_N_inventor	0.660	0.590	1.95	0
ln_N_Forward_Citation	0.481	0.661	2.56	0
US_Grant_dummy	0.069	0.254	1	0
subAppln_dummy	0.0292	0.168	1	0

Table III-2 Summary results of ordinary least squares (OLS) regression of the incidence of narrowing on the logarithm of the number of applicant citations

	Dependent Variable: *incidence of narrowing* (narrowed=1, unchanged=0)			
	(1)	(2)	(3)	(4)
	C & C	E & E	Mech.	Others
Independent variables	OLS	OLS	OLS	OLS
ln_N_Backward_Citation	-.00947***	-.0147***	-.00874***	-.00885***
	(.00272)	(.00203)	(.00240)	(.00330)
ln_inverse_claim_length(initial)	.208***	.173***	.197***	.206***
	(.00252)	(.00228)	(.00242)	(.00304)
ln_N_claims(initial)	.0453***	.0465***	.0514***	.0582***
	(.00189)	(.00177)	(.00190)	(.00255)
ln_N_inventor	.00625***	.00602***	.00276	.0140***
	(.00241)	(.00211)	(.00222)	(.00289)
ln_N_Forward_Citation	.0209***	.0182***	.0168***	.0279***
	(.00185)	(.00150)	(.00168)	(.00233)
US_Grant_dummy	-.000694	.0103***	.00663**	-.000717
	(.00298)	(.00278)	(.00323)	(.00640)
subAppln_dummy	.105***	.118***	.111***	.124***
	(.00547)	(.00552)	(.00679)	(.00889)
effective filing year	yes	yes	yes	yes
33_clasiffication	yes	yes	yes	yes
*effective filing year * 33_clasiffication*	yes	yes	yes	yes
applicant	FE	FE	FE	FE
Observations	115,889	142,831	138,010	90,708
R-squared	.0849	.0731	.0798	.0912
Number of applicant	1,268	1,746	2,209	2,230
adjusted R-Squared	0.0745	0.0615	0.0643	0.0671
Log Likelihood	-61825	-78243	-75272	-52169
Standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.10				

Table III-3 Descriptive statistics for the estimations reported in Table III-4

C & C	N=84,662			
variable	mean	sd	p99	p1
extent of narrowing	-6.80	1.21	-4.66	-10.8
ln_N_Backward_Citation	0.327	0.498	1.79	0
ln_N_claims(initial)	1.87	0.792	3.78	0
ln_N_inventor	0.483	0.585	1.95	0
ln_N_Forward_Citation	0.563	0.708	2.71	0
US_Grant_dummy	0.306	0.461	1	0
subAppln_dummy	0.0721	0.259	1	0

E & E	N=103,404			
variable	mean	sd	p99	p1
extent of narrowing	-6.71	1.21	-4.41	-10.5
ln_N_Backward_Citation	0.458	0.635	2.56	0
ln_N_claims(initial)	1.70	0.776	3.56	0
ln_N_inventor	0.624	0.605	1.95	0
ln_N_Forward_Citation	0.654	0.805	3.18	0
US_Grant_dummy	0.288	0.453	1	0
subAppln_dummy	0.0589	0.235	1	0

Mech.	N=99,589			
variable	mean	sd	p99	p1
extent of narrowing	-6.80	1.20	-4.53	-10.6
ln_N_Backward_Citation	0.423	0.545	2.08	0
ln_N_claims(initial)	1.42	0.741	3.18	0
ln_N_inventor	0.579	0.576	1.95	0
ln_N_Forward_Citation	0.578	0.726	2.83	0
US_Grant_dummy	0.202	0.402	1	0
subAppln_dummy	0.0385	0.192	1	0

Others	N=60,878			
variable	mean	sd	p99	p1
extent of narrowing	-6.53	1.22	-4.13	-10.3
ln_N_Backward_Citation	0.344	0.523	2.08	0
ln_N_claims(initial)	1.40	0.702	3.04	0
ln_N_inventor	0.683	0.596	1.95	0
ln_N_Forward_Citation	0.510	0.679	2.64	0
US_Grant_dummy	0.073	0.260	1	0
subAppln_dummy	0.0386	0.193	1	0

Table III-4 Summary results of ordinary least squares (OLS) regression of the extent of narrowing against the logarithm of the number of applicant citations

	Dependent Variable: *extent of narrowing*			
	ln(inverse claim length(initial) - inverse claim_length(grant))			
	(1)	(2)	(3)	(4)
	C & C	E & E	Mech.	Others
Independent variables	OLS	OLS	OLS	OLS
ln_N_Backward_Citation	-.0652***	-.0673***	-.126***	-.114***
	(.00905)	(.00659)	(.00763)	(.0104)
ln_N_claims(initial)	.157***	.188***	.218***	.224***
	(.00632)	(.00583)	(.00611)	(.00814)
ln_N_inventor	-.0172**	.0164**	-.00781	-.00908
	(.00789)	(.00684)	(.00705)	(.00904)
ln_N_Forward_Citation	.0275***	.0234***	.0166***	.0280***
	(.00599)	(.00480)	(.00529)	(.00719)
US_Grant_dummy	-.0784***	-.112***	-.133***	-.122***
	(.00984)	(.00899)	(.0101)	(.0198)
subAppln_dummy	.122***	.0973***	.161***	.0887***
	(.0164)	(.0162)	(.0194)	(.0246)
effective filing year	yes	yes	yes	yes
33_clasiffication	yes	yes	yes	yes
*effective filing year * 33_clasiffication*	yes	yes	yes	yes
applicant	FE	FE	FE	FE
Observations	84,662	103,404	99,589	60,878
R-squared	.0172	.0155	.0249	.0276
Number of applicant	1,139	1,567	2,048	1,984
adjusted R-Squared	0.00339	0.000104	0.00367	-0.00716
Log Likelihood	-133705	-162717	-153480	-92488
Standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.10				