Effects of early patent disclosure on knowledge dissemination: evidence from the pre-grant publication system introduced in the United States

Yoshimi Okada
Sadao Nagaoka

IIR Working Paper WP#15-12

Jul. 2015

Institute of Innovation Research
Hitotsubashi University
Tokyo, Japan
http://www.iir.hit-u.ac.jp

Effects of early patent disclosure on knowledge dissemination: evidence from the pre-grant publication system introduced in the United States

July 2015

Yoshimi Okada
Institute of Innovation Research, Hitotsubashi University

Sadao Nagaoka
Tokyo Keizai University

Abstract

In order to assess the disclosure function of the patent system, this study examined the impact of the pre-grant publication system introduced in the United States in 2000. Unlike earlier studies, the applicant (inventor) non-self-citations (excluding examiner citations) were used to track knowledge flow. The causal effects of disclosure were identified by examining the changes in behavior before and after this legal change. The introduction of the pre-grant publication system was found to accelerate the initiation of knowledge diffusion significantly across all technology areas, except for Chemical field. The effect was the strongest in the Computers & Communications field, which had the longest publication lag before the reform. In addition, the initial slope of the diffusion curve rose while the long-term level of citation flow declined in the Computers & Communications field. In contrast, both of them rose in the Electrical & Electronic field. These results suggest the possibility that early disclosure not only stimulated complementary inventions but also helped inventors recognize early the duplications and then helped the reductions of duplicative R&D and/or applications, in a field with a long publication lag. In addition, we found that the examiner citation curve begins significantly earlier and more sharply compared to the applicant citation curve, which shows that examiner citation is a wrong measure of knowledge flow.

Keywords: Disclosure, knowledge flow, patent, pre-grant publication
JEL: O33, O31, O38

1. Introduction and Research Objective

The patent system plays two roles or functions in promoting innovation. The first is to protect inventions from imitation for a certain period of time and thereby to promote investment in research and development (R&D) and commercialization. The second role is to disclose useful technological information to the public, facilitate the diffusion of technological knowledge, and avoid duplicated R&D activities. This study analyzes the second role, or more specifically, the effects of an 18-month pre-grant publication system on the dissemination of the technological knowledge disclosed in the patent application documents. The second function is fundamental to the patent system's contribution to the efficiency of R&D investments.

In order to examine the second role of the patent system, we analyzed the impacts of the introduction of an 18-month pre-grant publication system in the U.S. Most industrialized countries have included the 18-month pre-grant publication system in their patent system. Under the 18-month pre-grant publication system, technological knowledge pertaining to pending patent applications is made open to the public 18 months after the effective filing date of the patent application. We define "effective filing date" as the earliest priority date for the applications claiming priority and as the filing date for the applications claiming no preceding priority. The U.S. introduced the 18-month pre-grant publication system by enacting the American Inventor's Protection Act (AIPA) in 1999. Pending applications filed on or after November 29, 2000 are made open to the public 18 months after their effective filling date, with one exception. The applicants of the patent applications for inventions that are filed only in the U.S. and not in foreign countries can choose to keep their inventions secret before the patent grant. Henceforward, we call the patent applications for inventions that are filed only in the U.S. and not in foreign countries as "purely domestic applications." However, crucially for our research, Graham & Hedge (2012, 2015) revealed that over 80% of purely domestic patent applications opted for 18-month pre-grant disclosure rather than keeping their inventions secret until the patent was granted.

The empirical investigation of the effects of the pre-grant publication system on the diffusion of technological knowledge is important for two reasons: the acceleration of knowledge diffusion is not the inevitable outcome, and the introduction of pre-grant publication provides a significant opportunity for analyzing how disclosure affects the invention process. Pre-grant publication does not affect knowledge flow if the knowledge diffusion starts with the applications: the inventor may disclose details about the inventions in research workshops, trade shows, or through publications once his/her patent application is completed (or even before the application, during the grace period).

Moreover, it does not affect knowledge flow if the subsequent inventors rely primarily on granted patents as their knowledge sources, either because the patents are granted swiftly enough for the inventors' decisions or because the information disclosed in the granted patents are regarded to be more trustworthy. Thus, whether pre-grant publication accelerates knowledge flow is an empirical issue.

While Johnson & Popp (2001) showed that the (total) citation flow begins with the publication date (or equivalently, the grant date), their study has a fundamental problem because of the inclusion of examiner citations as the measures of knowledge flow. Examiner citations make us observe exactly such correlations, even without the knowledge diffusion effect. Examiner citations begin with the publications because examiners can cite only published documents; simultaneously, they are required to undertake an ex-post search of all documents that are public at the priority date of the examined applications but about which the inventors may not know at the time of their inventions (section 5). Additionally, their results are based on the variations in grant lag, which may make a causal interpretation more difficult. Our study aims to identify the knowledge diffusion effects of disclosure based on the changes in the applicants' non-self-citations in response to the new disclosure rule under the AIPA. We use purely domestic applications for our study since these publications are the first and only disclosure of the inventions to the public; thus, their pre-grant publications constitute new disclosure under the AIPA.

Additionally, this study examines the effects of the pre-grant publication system on the pattern of knowledge diffusion, in order to explore the relative importance of the two potential roles of disclosure: reducing duplications in R&D and/or patenting and accelerating the increase of knowledge stock for complementary inventions. In this regard, it is important to recognize that there are two types of knowledge (citation) flows: those to the follow-up inventions based on the focal inventions and those to the independently made duplicative inventions, and that they have significantly different citation flows from each other. The latter citation flows emerge in the period immediately following the publication of the focal cited patent since the US patent law requires the inventors to disclose not only the prior arts they know when they invented but also those they came to know by the time of examination of their applications. Moreover, the duplicative inventions will decline over time as the inventors will avoid the duplicative R&D and patent applications with respect to the disclosed inventions. Thus, if early disclosure results in an immediate increase of the slope of the diffusion curve but in a reduction in the long-term level of citation flow from the focal patent, it would be consistent with the existence of duplicative R&D and/or patent applications and with their subsequent reductions caused by early disclosure. On the other hand, if early disclosure results in an increase of both the long-term level of citation flow as well as of its initial slope, it would be consistent with the accelerated increase of knowledge stock for complementary inventions. We expect the duplication reduction effect of early disclosure would be stronger in fields where the disclosure was late prior to the legal change. To the best of our knowledge, such effects have not been examined by the existing literature. We find that the long-term level of citation flow declined significantly only in the Computers & Communications field, where the initial slope increased highly significantly. The publications lag before the reform was very long (around 48 months) in this field.

The rest of this paper is organized as follows. Section 2 presents a brief review of the extant literature. Section 3 describes the institutional background and presents the hypotheses. Section 4 explains the data construction and presents the econometric model. Section 5 presents the estimation results and discussions, and section 6 concludes the paper with directions for future research.

2. Literature Review

Graham & Hedge (2012, 2015) found that the applicants of purely domestic patent applications often opted for 18-month pre-grant disclosure rather than keeping their inventions secret before the patent grants (by not opting for 18-month pre-grant publications). This finding is important for our study because the U.S. reform in 2000 could have had a genuine impact on the acceleration of disclosure, even though purely domestic applications are exempt from compulsory pre-grant publications. Further, using (i) the time lag between application and grant, (ii) the number of claims in a patent, (iii) patent renewal rates, and (iv) the number of later granted U.S. patents that cite the focal patent as prior art, Graham & Hedge (2012, 2015) found that small applicants chose disclosures over secrecy for important inventions.

Building on the model developed by Caballero & Jaffe (1993), Johnson & Popp (2001) investigated patent citation data between 1976 and 1999 and showed that the diffusion process begins with the publication (or the grant) of the patent (not with the application). They concluded that introducing an 18-month pre-grant publication system would facilitate the diffusion of information by reducing the delay between filing and publication. However, they did not differentiate examiner citations and applicant (inventor) citations, or self-citations and citations by others, which is the fundamental problem of their study. Examiner citations and self-citations do not represent knowledge flow to the inventors through patent disclosures. If we use the entire citation data in the analysis, we are bound to observe strong but spurious correlations between the introduction of pre-grant publications and the acceleration of citation flows, even when there are no such effects of pre-grant publications, for the following reasons. First, the examiner citation flows associated with a particular patent application begin with its earliest (pre-grant, if it exists) publication because examiners can cite only published documents, and they are required to search ex-post (at the time of patent examination, not at the time of patent application) all the documents that are public at the priority date of the examined applications.[1] Therefore, examiners cite all published applications that are relevant to the examination of the focal patent application, including those about which the inventor of the focal application may have no knowledge at all before his/her application. Because of this process, the examiner citation flows inevitably begin with the first publication of the patent application. Moreover, these examiner citations account for about half of the citations. While it has been well-known since the earliest years of exploiting the patent citations in economic analysis that citations are often added by those other than inventors, it is not a random noise[2] for the purpose of analyzing the effect of pre-grant publications on knowledge diffusion. Rather, it is the cause of a systematic bias (in section 5, we show that examiner citations begin much more sharply with the publication than inventor citations do).

The second problem of Johnson & Popp's (2001) study is that their results related to the effect of publication on knowledge flow are based on the variations in grant lags; this may make a causal interpretation difficult because of the endogeneity of grant lags. The citation flows may begin slowly for a patent with a long grant lag not because the publication is delayed but because such a patent does not have much knowledge source for further diffusion. Therefore, our study examines the changes in knowledge flow accompanying the introduction of the pre-grant publication system. We compare the applicant (inventor) citations, excluding self-citations, before and after the legal change.

Hegde & Luo (2013) investigated the effects of pre-grant publications on patent licenses in biomedical technology to examine how the disclosure of a patent application facilitates the licensing, separately from the effect of the grant.[3] They found that post-AIPA patent applications are licensed significantly sooner than pre-AIPA patent

---

[1] See Alcacer & Gittelman (2004, 2006) and Alcacer, Gittelman, & Sampat (2009) for detailed explanations on how examiner and applicant citations are generated and for general discussions about their implications in the analysis of knowledge flow and patent value.

[2] See Jaffe, Trajtenberg & Henderson (1993) and especially Jaffe, Fogarty & Banks (1998) for the "noise" of patent citations as a measure of knowledge flow.

[3] Gans, Hsu & Stern (2008) discussed the importance of the certification of inventors' property rights through patent grants for the licensing contract.

applications controlling for the date of grants, especially those that do not have equivalent foreign applications. More specifically, they found the following: i) the likelihood of licensing in the window between 18-month publication and patent allowance more than doubles for post-AIPA patent applications; ii) post-AIPA patent applications are about 18 percentage points less likely to wait until allowance to be licensed; iii) the overall effects of AIPA are stronger for U.S. patent applications that do not have equivalent foreign applications. Thus, they concluded that pre-grant patent publications appear to facilitate transactions in the market for ideas and significantly accelerate the commercialization of inventions. While this prior study convincingly showed that disclosure facilitates licensing, it did not show the effect of early disclosure on knowledge flow with regard to subsequent inventions.

The extant studies have not analyzed the relative importance of the two potential roles of disclosure: avoiding duplications in R&D and/or patenting and accelerating the increase of knowledge stock for complementary inventions. In this context, it is important to note the difference between "fewer duplications" and "more obsolescence," although we observe a decline in follow-up inventions in both cases. The obsolescence of the focal invention occurs because of the arrival of inventions competing with the focal invention, while the duplications of the focal invention are reduced by its early disclosure, quite independently from the arrival of competing inventions.

3. Institutional Background and Hypotheses on Knowledge Flow

(1) Changes caused by introduction of 18-month pre-grant publication system

The introduction of the 18-month pre-grant publication system affects knowledge dissemination through two different types of changes: the acceleration of the timing of publication and the expansion of the coverage of patent applications to be made public.

Figure 1 shows the change in the timing of publication before and after the patent law reform. Before the AIPA came into force, patent applications remained secret until the patent issue date. In other words, the publication and the patent grant were simultaneous events. After the AIPA came into force, pending patent applications are published 18 months after the filing date, which is usually earlier than the grant date. We define "publication lag" as "the difference between the filing date and the first publication date" and "grant lag" as "the difference between the filing date and the grant date." The publication lag is reduced by the introduction of the pre-grant publication system, which is expected to accelerate and enhance the diffusion of knowledge flow from the patent applications.

(Figure 1)

In Figure 1, we considered only U.S. domestic applications. However, in many cases, applicants file patent applications for the same invention in foreign countries. Figure 2 illustrates this case. If a U.S. applicant files for a patent in a foreign country, the corresponding foreign equivalent patent application is published 18 months after its priority date. Thus, the contents of the original U.S. patent application are effectively made open to the public through the pre-grant publication of the foreign equivalent patent application(s), although the publication language might not be English, which might pose a language barrier for U.S. nationals. Even before the AIPA law reform, knowledge diffusion could begin through the pre-grant publications of the foreign equivalent patent applications. In contrast, the contents of a purely domestic patent application that is filed only in the U.S. are made open to the public only through the publication made by the United States Patent and Trademark Office (USPTO). Since the effect of pre-grant publications made by foreign governments is not due to the AIPA law change, we focus on "purely domestic applications" for our analysis.

(Figure 2)

Table 1 presents the statistics on publication and grant lags for the purely domestic applications that claimed no priority based on earlier applications[4] and that were eventually granted patents, which are the sample of our econometric analysis. The mean publication lag (A), the mean grant lag (B), the difference (B - A), the standard deviation of the publication lag, and the number of granted publications are indicated by technological fields for each 12-month period 3 years before and after the law reform. The PATSTAT (2013 October edition) is used as our database. The average reduction in publication lag for the applications filed between the 1-year period up to December 2000 and the 1-year period up to November 2001 amounted to around 20 months in the Computers & Communications field (the largest among the six technology fields considered in this study); the average reduction was several months in other fields (5 months for the Mechanical field and 7.6 months for the Drugs & Medical field). [5]Since

[4] Applications claiming priority based on provisional applications and secondary applications such as continuation applications, divisional applications, and continuation-in-part applications are not included in this analysis. Henceforward, we call such applications "normal applications."

[5] We can estimate the ratio of opting for pre-grant publication by calculating "the number of granted patents of which pre-grant publications were made" divided by "the number of granted patents." According to the data of Figure 4, the ratio for the patents filed in 2001 in each technological fields are as follows: Chemical, 61%; Computers & Communications, 67%; Drugs & Medical, 63%; Electrical & Electronic, 55%; Mechanical, 61%; and Others, 56%. The ratio is largest in Computers & Communications field, and second largest in Drugs & Medical field. According to table 1, mean grant lag in patents filed between December 2000 and November 2001 is largest in Computers &

the applications filed before November 29, 2000 could opt for pre-grant disclosures only by taking special procedures, they were exceptions. Thus, there are only very small differences between the publication lags and the grant lags before the AIPA law reform.

(Table 1)

Figure 3 indicates the change in the coverage of the patent applications to be disclosed, which was caused by the introduction of the pre-grant publication system. Before the AIPA reform, applications that were rejected and were eventually deemed to be abandoned were not published. After the AIPA reform, the pending applications that are eventually rejected and deemed to be abandoned are published as well. Since these publications are also potential sources of knowledge flows, we can expect an increase in the knowledge diffusion flow through the expansion in the coverage of the patent applications subject to disclosure.

(Figure 3)

Figure 4 presents the number and the composition of publications of purely domestic normal applications over the filing years from 1990 to 2011. Publications are classified into three categories: i) patent publications without pre-grant publications (indicated in blue); ii) patent publications for which pre-grant publications were published (indicated in red); and iii) pre-grant publications that are not yet granted patents (indicated in green). Because the third category (green) includes abandoned applications as well as pending applications, its share becomes larger in subsequent filing years. According to Carley, Hegde & Marco (2013), the influence of pending applications is negligible for patent applications that were filed before 2002. They reported that the grant rate including requests for continued examination (RCE) after the final rejection (the application serial number did not change by this procedure) is about 70% in 2001, which is consistent with our data (the ratio of red and green is about 2:1).

(Figure 4)

(2) Applicant (inventor) citation data to measure knowledge flow

We use applicant (inventor) citation data to measure knowledge flow. When a patent application is filed to the USPTO, each individual responsible for the application (the applicant, the inventors, and the agent) has a duty to disclose to the USPTO all information known to that individual to be material to its patentability. In order to fulfill this duty, applicants usually submit the Information Disclosure Statement (IDS).

Communications, 51 months and second largest in Drugs & Medical, 31 months. This is why the publication lag is largest in Computers & Communications and second largest in Drugs & Medical.

We believe that the IDS citation data effectively reflect what applicants (inventors) actually know. Thus, applicant citation is a good measure of knowledge flow for the following legal reasons.

There are two legal effects of submitting the IDS. First, the risk of unenforceability in future litigation is diminished (Cotropia, Lemley & Sampat, 2013). If an applicant deliberately has not disclosed a prior art that is material to patentability, an accused infringer can raise an issue of "inequitable conduct" that renders the patent unenforceable. By submitting the IDS, the applicant can diminish this risk. The second effect is to allow the applicant to acquire the strengthened presumption of validity (Juneau & MacAlpine, 2000; Buchanan, 2006; Allison & Lemley, 1998; KSR International, Co. v. Teleflex, Inc., 2007). Presumption of validity against the group of prior arts that were considered by the patent examiner is stronger compared to that against the prior arts that were not considered by the patent examiner. Because of these legal effects,[6] there is a strong incentive and a legal mechanism (i.e., the IDS) for the applicants to submit all the prior art documents that they are aware of to the USPTO.

(3) Availability of U.S. applicant citation data in PATSTAT

We used the PATSTAT database (2013 October edition) provided by the European Patent Office for the citation data of U.S. patents that originated in the reference data of U.S. patents. There are two origins of the reference citations associated with U.S. granted patents: "Notice of References Cited" (form PTO-892) and the IDS. The former is prepared by the patent examiner during the patent examination process, and the latter is provided by the patent applicant. Since 2001, the references listed by the patent examiner in a "Notice of References Cited" are indicated with an asterisk in the "References Cited" section on the front page of a patent document. Regarding the "Notice of References Cited," the Manual of Patent Examination Procedure (MPEP section 2001) states that "the examiner does not list references which were previously cited by the applicant (and initialed by an examiner) on an Information Disclosure Statement." Therefore, the overlap between the references cited in the "Notice of References Cited" (i.e., examiner citation) and the references disclosed in the IDS is expected to be very small. Citation origin data are available only where the publication year of the citing patents is in and after 2001.

We investigated randomly chosen 50 sample patents issued between 2005 and 2009 by examining the associated image file wrapper using the public PAIR system.

---

[6] In addition to these two effects, there are strong incentives for the patent attorney to fulfill the duty. If a patent attorney deliberately ignores the duty, he/she might be in danger of disqualification.

Figure 5 illustrates the results. The overlap between PTO-892 cited documents and IDS disclosed documents is very small, and 94% of the IDS disclosed documents are reflected in the applicant citation data, as indicated in the "References Cited" section of U.S. patent documents. By using the patent citation data in the PATSTAT database, we can effectively capture the references that are actually disclosed by the applicant in the IDS. In contrast, if we use the entire citation data, 42% of the cited documents are cited by only the patent examiner, and the noise is quite large.

(Figure 5)

(4) Comparing applicant citation flows and examiner citation flows

Figure 6 presents the citation-lag dependence of probability of both applicant citation and examiner citation in the Computers & Communications and Electrical & Electronic fields (We explain construction of these data in detail later in section 4.4.). The shape of the probability of citation vs. citation-lag curve is quite different between the examiner and applicant citations. The examiner citation curve begins significantly earlier and more sharply than the applicant citation curve in both the Computer & Communications and Electrical & Electronic fields. The peak of the examiner citation curve is around the 18-month citation lag, and it comes much earlier than that of the applicant citation curve. If we use total citations as the measure of knowledge flow, we would get superfluous results about knowledge flow. Thus, it is essential to use applicant citation data to measure knowledge flow. We will discuss this later in detail in section 5.2.

(Figure 6)

(5) Hypotheses on the effects of early disclosure on knowledge flow

In analyzing the effects of early disclosure, we consider two types of knowledge flows. First, disclosure generates follow-up inventions, until the focal invention becomes obsolete stochastically. For simplicity, we assume no obsolescence of the focal patent in the rest of this section (we do take this into account in our empirical analysis). Second, disclosure can affect the patentability and the patent scope of the duplicative inventions that are concurrently made by other firms. In this case, disclosure does not generate new inventions but affects the scope of the patents of independently made inventions. In both cases, the citations are made with reference to the focal patent; thus, the citation flows reflects these two types of knowledge flows.

There is a major difference in the pattern of the diffusion curve following the focal invention of these two types of knowledge flows. We expect that the first type of knowledge flow starts gradually from the level of zero given the time necessary for the recognition and invention, increases over time, and subsequently levels off (unless there is obsolescence), as assumed by Caballero & Jaffe (1993) and Johnson & Popp (2001). On the other hand, the second type of knowledge flow increases rapidly initially, since the citing inventions are independently made and not endogenously developed based on the focal invention, however, it subsequently declines toward zero since the firms avoid duplicative inventions.

We assume that the combined knowledge flow can be approximated by the following logistic function in equation (1), where T represents the time, t is the date of the patent application of the focal invention, and τ gives the lag between the application date and the inflection point[7]:

$$x(T;t,\tau) = \frac{\alpha}{1+e^{-\mu(T-t-\tau)}} \qquad (1)$$

In equation (1), the parameter for the ceiling level of citation flow (α) measures the long-term level of the inventions generated by the focal patent; it is affected only by the first type of knowledge flow. The parameter (μ) indicates how rapidly the knowledge flow increases for a small x (the slope of the diffusion curve); both types of knowledge diffusions affect this parameter.

Next, we consider the effects of early disclosure. The direct effect would be an acceleration of the initiation of diffusion from the date of the patent application of the focal patent (t), if the pre-grant publication results in early disclosure. We can measure this by the early arrival of the inflection point (smaller τ).

In addition, early disclosure would expand the total knowledge stock based on the focal patent that is available for an inventor for a new invention at each point of time, since the inventions based on the focal patent are also disclosed early. Thus, when these disclosed inventions are all complementary as knowledge stock for creating new inventions, early disclosure will accelerate the growth of the knowledge stock available for new inventions. This effect will result in both the increase in the long-term level of citation flow (α) as well as the increase of the initial slope parameter (μ) in this model .

Moreover, when the duplicative R&D and patent applications are important, early disclosure can help an inventor recognize early the focal invention and avoid duplicative R&D investments and/or patent applications in the long run. The early disclosure will

[7] This would be the case if the second type of knowledge spillover is relatively small.

force the inventors pursuing similar R&D projects to cite the focal patent disclosed earlier when applying for patent protections for their inventions. If an inventor applies for a patent for a duplicative invention without knowing about the focal patent, the inventor will not (cannot) cite the focal patent at the time of patent application. On the other hand, if the focal patent is already published and the inventor of a duplicative invention recognizes it, the focal patent has to be cited. Thus, the early disclosure of the focal patent will result in more citation flows via these duplicative inventions immediately after its pre-grant publication, resulting in a significant increase of the initial slope parameter (μ). At the same time, the number of granted patents citing the disclosed invention will decline in the long run because of the early disclosure. That is, there will be a reduction in the long-term level of citation flow (α) in this model.

The preceding discussions are summarized in the following three hypotheses. The first hypothesis to be tested in our study is related to the timing of the initiation of knowledge diffusion, counted from the date of the patent application of the focal patent. It is essentially the same hypothesis as tested by Johnson & Popp (2001). However, we aim at providing evidence allowing a more causal interpretation by using actual data based on applicant non-self-citations that are generated after the legal change.

> *Hypothesis 1 on earlier initiation of knowledge diffusion:*
> The introduction of the pre-grant publication system accelerated the initiation of knowledge diffusion, especially in technological fields where the grant lag was long.

The second hypothesis is related to the heterogeneous impact of early disclosure on the shape of knowledge curve, depending on the relative importance of citation flows indicating cumulative inventions and those indicating the duplicative inventions

> *Hypothesis 2 on fewer duplications vs. larger knowledge stock*
> The effects of pre-grant publications on the shape of the diffusion curve can be heterogeneous, depending on the relative importance of the two citation flows. If early disclosure results in an increase of the initial slope of the diffusion curve but in a reduction of the long-term level of citation flow from the focal patent, it would be consistent with the existence of duplicative R&D and/or patent applications and with their subsequent reductions by early disclosure. On the other hand, if early disclosure results in an increase of both the long-term level

of citation flow as well as of its initial slope, it would be consistent with the accelerated increase of knowledge stock for complementary inventions.

Additionally, the pre-grant publication system would have expanded the disclosure of those inventions that are not granted patents.

Hypothesis 3 on knowledge flow effect of more disclosures
The pre-grant publication system increased the sources of knowledge flows since the pending patent applications that were eventually abandoned and not granted patents were also published under this system.

4 . Data and Econometric Model

(1) Constructing data for non-self-citations by applicants

In order to construct the applicant citation data excluding self-citations, we first constructed a citing-cited matrix based on the U.S. patent bibliographic data for both citing and cited patent documents using the PATSTAT database (2013 October edition). This covers the effective filing date, patent classification, and citation origins (whether the references are provided by the applicants or the examiners).

In order to construct samples that are comparable with regard to the average quality of the inventions before and after the introduction of the pre-grant publication system, we selected the publications of granted patents and the pre-grant publications that were eventually granted patents as our dataset of cited patent documents. Cited patent applications were limited to purely domestic patent applications that were filed only in the U.S. and not in foreign countries. We think that the introduction of pre-grant publication does not significantly affect the choice of the applicants for purely domestic patent applications. After the legal reform, applicants of purely domestic applications can still choose keeping their inventions secret before the patent grant. Therefore, introduction of pre-grant publication system does not work as a deterrent for the applicants that opt for secrecy before grant. Our cited patent data do include the patents of which applicants opted for pre-grant secrecy. Moreover, the introduction of pre-grant publication system does not affect the behavior of the applicants who opt for filing overseas. Such applicants need to apply for foreign applications irrespective of the introduction of the pre-grant publications.

Note that the applicants are not restricted to U.S. residents. The data included applications filed by foreign applicants seeking U.S. patents only (and no foreign patents) for an invention. Secondary applications or progenitor applications (such as continuation applications, divisional applications, and continuation-in-part applications) were removed from the dataset of both citing and cited applications. One reason for this decision was to measure accurately the timing of when the invention was disclosed to the public. For example, the knowledge diffusion might have begun with the publication of the preceding original parent patent application rather than with that of the focal secondary or progenitor patent applications, if the original application was published and contained the same information as that of the focal patent application. By removing the secondary (progenitor) applications from our dataset of cited patent documents, we did not have to consider this type of knowledge diffusion. The second reason was to reduce ex-post citations. Since the IDS obligation continues until the patent issue date, the IDS includes not only the applicants' knowledge at the filing date (ex-ante knowledge) but also the knowledge that the applicants or their agent acquired after the filing date (ex-post knowledge). If we include secondary (progenitor) applications in the dataset of citing applications, the extent of ex-post knowledge contamination would increase significantly. For example, applicants of secondary (progenitor) applications usually include in the IDS those references that were provided by the patent examiner during the original parent application process. In order to measure the timing of knowledge flow accurately, we dropped those applications that involved priority based on provisional application(s) from the dataset of cited patent documents. Therefore, the effective filing date of the cited patent is the same as the filing date.

In order to distinguish non-self-citation from self-citation in the citation matrix data, we used the patent assignee data prepared under the NBER Patent Data Project (https://sites.google.com/site/patentdataproject/). The data provide the patent assignee-identified data for U.S. Patents that were published between 1976 and 2006. We prepared the concordance table between the patent assignee identification numbers of PATSTAT and the patent assignee identification numbers prepared under the NBER Patent Data Project using the granted patent numbers of single patent owners as the connecting keys. Using the data in this concordance table, the PATSTAT applicant ID data in the citation matrix is transformed to the NBER assignee ID data. We restricted cited publications to those publications whose assignees were identified by the NBER assignee ID data. Henceforward, we call these applications "applications by the NBER-identified assignees." We classified a citation as a self-citation if the NBER assignee ID was the same for the cited and the citing patent document. For jointly-held patents, the citations were classified as self-citation if at least one owner was common in the cited and the citing patents. Otherwise, we classified the citations as non-self-citation. Finally, self-citations were removed from the citation matrix.

Table 2 provides the descriptive statistics on the publication and grant lags for the granted purely domestic normal applications by the NBER-identified assignees. A comparison of Table 2 and Table 1 suggests that there is no major sample selection bias in terms of publication lag and grant lag because of our focus on the patents of NBER-identified assignees.

(Table 2)

We divided the applications into six technological categories following Hall, Jaffe, & Trajtenberg (2001): "Chemical," "Computers & Communications," "Drugs & Medical," "Electrical & Electronic," "Mechanical," and "Others." The data included several classification codes that were not listed in Hall, Jaffe, & Trajtenberg (2001). In those cases, we classified them into an appropriate technological category by referring to the category and subcategory assignment list prepared by the NBER patent project (https://sites.google.com/site/patentdataproject/Home/downloads/patn-data-description). The data pertaining to the United States Patent Classification is missing in publications that were published in and after September 2012 in the PATSTAT database (2013 October edition). This is another cause of truncations of the citation data, in addition to the pending patent examination status of potentially citing patent applications.

(2) Constructing data of Citation probability among cohorts

We divided the citation matrix data (section 4.2) into cohorts ($C_{co,i,t,I,T}$) according to citation origin (*co*), cited patent's technological field (*i*), cited patent's effective filing time measured by calendar month (*t*), citing patent's technological field (*I*), and citing patent's effective filing time measured by calendar month (*T*). Additionally, we measured the size of potentially citing cohorts by the number of granted patents ($M_{I,T}$) and the size of potentially cited patent document cohorts by the number of purely domestic normal patent applications by the NBER-identified assignees ($N_{i,t}$). Note that $M_{I,T}$ is not restricted to the publication of purely domestic applications. $M_{I,T}$ includes the patents that were filed internationally and the patents claiming priority based on provisional applications.

Next, we obtained the probability of citation between each pair of cohorts (the cited patent's filing time t and the citing patent's filing time T) for each technical field ($i=I$): $P_{co,i,t,I,T}$, using the following formula:

$$P_{co,i,t,I,T} = \frac{C_{co,i,t,I,T}}{N_{i,t} * M_{I,T}} \quad (3)$$

(3) Econometric model

To analyze the dynamics of the patent citation flows, we used the model in Equation (4), which is similar to what Johnson & Popp (2001) used. The dependent variable is the probability of citations between a pair of cohorts:

$$P_{co,i,t,I,T} = \frac{\alpha}{1+ e^{-\mu*(T-t-\tau)}} * e^{-\lambda*\Sigma N_i} \quad (4)$$

In Equation (4), ΣNi represents the total number of patent documents that are published between the cited patent's filing time in month t and the citing patent's filing time in month T in technological category i. While calculating ΣN_i, only the first publication for each application is counted in order to eliminate duplication in the count of the publications that originated from the same invention. Patent publications categorized as "B2" (which were published after the pre-grant publications) were not counted. Publications for secondary (progenitor) applications (such as continuation applications, divisional applications, and continuation-in-part applications) were not counted because all or at least the large part of the contents of these applications would have been disclosed in the preceding parent applications.

The model consists of two factors: i) a diffusion factor with the ceiling level constant (α), and ii) an obsolescence or decay factor (new inventions make older patents obsolete). We modified Johnson & Popp's (2001) model in the following four aspects. (1) While Johnson & Popp (2001) used the difference in publication timing between citing and cited patents as citation lag, we used the difference in effective filing date between citing and cited patent documents as citation lag. This is because we want to measure the effects of early publications on the initiation of knowledge diffusion counted from the date of application. (2) While Johnson & Popp (2001) used the number of patents in all technological fields as a determinant of the decay factor, we used the number of patent documents in the field, which is the same as that of the cited patent documents. We believe this approach makes more sense, given the heterogeneity of the pace of technological progress across technology fields. (3) We focused on the probability of citations where the citing patent document and cited patent document belong to the same technological category. (4) While Johnson & Popp (2001) used *(1-exp(-μ * citation_lag))* as the diffusion model, we used the logistic function (α*/(1+exp(-μ * citation_lag))*), following the standard used in the diffusion literature. In our regression analysis, the former function performed much worse than the latter did. The main reason for the differential performance of these two models seems to be the fact that Johnson & Popp (2001) used the total citations, more than 40% of which are examiner citations, and they used the difference in publication timing between citing and cited patents as the citation lag. We believe that our model is better given the more appropriate choice of citation data for measuring knowledge flows.

In order to investigate the changes in knowledge diffusion flow caused by the legal revision, we used a dummy variable representing the date of the introduction of the pre-grant publication system (*rev*), which is set to 0 if the filing date of the cited patent is between December 1999 and November 2000 and is set to 1 if the filing date of the cited patent is between December 2000 and November 2001. We used three models (Equations (5) to (7)). Coefficient α_1 is a ceiling level of citation probability in the logistic function of the information diffusion factor. Coefficient μ_1 stands for the slope parameter in the logistic function of the information diffusion factor. Coefficient τ_1 represents the location of the inflection point of the logistic function, with the horizontal axis representing the citation lag *(T-t)*, where the value rises most steeply. Coefficient λ stands for the decay parameter of the obsolescence factor in the focal technological field. Coefficient α_2 measures the change in the level of the ceiling of the citation probability caused by the law revision. Coefficient τ_2 measures the acceleration in the diffusion caused by the law revision. Coefficient μ_2 measures the increase of the slope of the diffusion curve caused by the law revision. Model 1 (Equation 5) is the most general base model for hypothesis testing, and it includes α_2, τ_2, and μ_2. Model 2 (Equation 6) and Model 3 (Equation 7) are simplified models. Model 2 includes only τ_2, and Model 3 includes only α_2.

Model 1 (the general base model for hypothesis testing)

$$P_{co,i,t,i,T} = (\alpha_1 + rev * \alpha_2) * \frac{1}{1+ e^{-(\mu_1+rev*\mu_2)*(T-t-(\tau_1+ rev * \tau_2))}} * e^{-\lambda*\Sigma N_i} \tag{5}$$

Model 2 (the model accommodating only the inflection point change)

$$P_{co,i,t,i,T} = \alpha_1 * \frac{1}{1+ e^{-\mu_1*(T-t-(\tau_1+ rev * \tau_2))}} * e^{-\lambda*\Sigma N_i} \tag{6}$$

Model 3 (the model accommodating only the ceiling level change)

$$P_{co,i,t,i,T} = (\alpha_1 + rev * \alpha_2) * \frac{1}{1+ e^{-\mu_1*(T-t-\tau_1)}} * e^{-\lambda*\Sigma N_i} \tag{7}$$

Based on the estimation results from Model 1, we can test the first two hypotheses. Hypothesis 1 implies that the pre-grant publication system causes the inflection point to arrive earlier (the estimated value of τ_2 is negative). Hypothesis 2 implies that the introduction of the pre-grant publication system makes μ_2 significantly positive and the ceiling α_2 negative if the initial duplications are substantive and the pre-grant publication has a significant effect of reducing such duplications. It also states that the introduction of the pre-grant publication system makes both α_2 and μ_2 positive if the pre-grant publication enhances complementary inventions.

Models 2 and 3 concentrate the total effects into one parameter, which provide a summary measure of the effects of pre-grant publication on knowledge diffusion. Model 2 provides the effective overall acceleration of the initiation of the knowledge diffusion, while Model 3 provides a summary of the change in the overall level of citation probability.

(4) Estimation Method

We implemented regression analysis of the applicant citation flow by nonlinear regression using Model 1 to Model 3 (Section 4.4). We used the data with the filing date of the cited patent documents before and after 1 year from the law revision, i.e., from December 1999 to November 2001. In order to reduce the influence of the truncation of citing patents, regression was performed in the range of 0 < citation lag <= 96 (month). To control for heteroskedasticity, we used analytical weights $(M_{I,T} * N_{i,t})$ with reporting heteroskedasticity robust standard error. Table 3 shows the descriptive statistics of the data used for the nonlinear regression.

(Table 3)

5. Results and Discussion

(1) Effects on knowledge diffusion curves

Table 4 presents a summary of the results for the six technology fields. Focusing on the field of Computers & Communications, where the publication lag was the longest before the introduction of pre-grant publication, Figure 6 shows the probability of actual citation and the fitted value over time using Model 1, comparing the citation probabilities for the cited patent documents with filing date between December 1999 to November 2000 and those with filing date between December 2000 to November 2001.

(Table 4)

In the Computers & Communications field, the legal change caused the inflection point to arrive significantly earlier (the estimated value of τ_2 is -16.0 months) according to Model 1. Thus, the results show a significant acceleration of knowledge flow, which is consistent with Hypothesis 1. This result is consistent with the significant decline (around 20 months) of the publication lag (the difference between the publication date and the effective filing date) because of the reform in this field.

Simultaneously, the long-term level of citation probability measured by α declined significantly (around 15%), while the slope of the diffusion curve measured by μ was enhanced significantly (by over 34%). These results are highly consistent with the possibility of an early finding of the existence of the close prior patent applications due to their early publications and the reduction of duplicative inventions in the long run. As stated earlier, the Computers & Communications field had the longest publication lag prior to the legal reform. Most observers would agree that it is a highly competitive technology field. The long publication lag in a highly competitive field might have made duplicative R&D activities and patent applications serious. These results support the duplication reduction effect of early disclosure in this field, as stated in Hypothesis 2.

As for the result from the simplified Model 2 that includes only τ_2, the legal change caused the inflection point to arrive significantly earlier overall (the estimated value of τ_2 is -8.87 months), which is actually smaller than the estimate obtained from Model 1. According to Model 3, which includes only α_2, a significant positive value of around 19% relative to α_1 was estimated. This means that the level of citation probability increased by 19% overall because of the law revision (note, however, that a substantial part of this increase is likely to reflect the citation flows indicating duplicative R&D and/or patenting in this field), even though the long-term level of citation flow declined.

(Figure 7)

Further, we observe significantly negative values of τ_2 (earlier start of diffusion) according to both Model 1 and Model 2 in all the other technological fields (except the Chemical field). The estimated length of the earlier start of diffusion varies from 3 months (Mechanical) to 6 months (Drugs & Medical) in Model 1. It varies from 4 months (Mechanical) to 9 months (Drugs & Medical) in Model 2. Thus, Hypothesis 1 is supported for four more fields: Electrical & Electronic, Drugs & Medical, Mechanical, and Others, in addition to Computers & Communications.

In the Chemical field, according to Model 1, we observe negative value of τ_2 at the 10 % significance level. However, in Model 2, we do not find statistically significance effect. Hypothesis 1 is marginally supported only in Model 1 and not in Model 2. There are many zero value points (more than 300) in the citation probability data. This means that the citation pair data is too small for obtaining strong and robust statistically significant results.

Moreover, we observe significantly positive coefficients for α_2 in Model 3 in these four fields. The estimated increase of the average long-term citation probability varies from 9% (Others) to 25% (Drugs & Medical) among the four fields with significant results. In this model, the effect of pre-grant publication is strongest in Drugs & Medical among the four fields. The decline in the publication lag is the largest in this field, although it is much smaller than that in the Computers & Communications field.

According to Model 1, in the Drugs & Medical and Electrical & Electronic fields, we observed significantly positive coefficients for α_2 and significantly negative coefficients for τ_2, while μ_2 is statistically insignificant for the Drugs & Medical field but significantly positive for the Electrical & Electronic field. These results are consistent with the greater number of complementary R&D projects in the two fields. Compared to the Computers & Communications field, the effect of facilitating complementary R&D is more dominant than the duplication reduction effect.

In summary, knowledge diffusion began earlier and was significantly facilitated by the introduction of the pre-grant publication system. Except in the Chemical field, there is significant evidence for the significant effects of pre-grant publications on the accelerated initiation of diffusion (five out of six fields). In addition, the initial slope of the diffusion curve rose while the long-term level of citation flow declined in the Computers & Communications field, which had the longest publication lag before the reform. In contrast, both of them rose in the Electrical & Electronic fields. These results suggest the possibility that early disclosure not only stimulated complementary inventions but also helped inventors recognize early the duplications and then helped their reductions, in a field with a long publication lag.

(2) Robustness checks

We did the following two robustness checks. First, we assessed the effect of ex-post knowledge on citation flow. The applicant's duty to candor (i.e., IDS obligation) continues until the patent is issued. Strictly speaking, the IDS contains not only the applicant's knowledge before filing (ex-ante knowledge) but also what he/she gets to know after filing the patent application (ex-post knowledge). Table 5 presents the number of citations (column A) where "publication lag > citation lag" and the number of citations (column B) where "publication lag <= citation lag." The former case involves ex-post knowledge because the applicants of citing applications cannot identify the cited applications made by the other firms that are not yet published. The share of case A was larger before the law revision compared to that after the revision (i.e., more ex-post knowledge was incorporated in citations before the introduction of the pre-grant publication system). This finding shows that the knowledge diffusion effects we found are robust when considering the inclusion of ex-post knowledge in the IDS information, which was larger before the legal change.

(Table 5)

As another robustness check on the effect related to the acceleration of diffusion, we directly assessed which of the pre-grant publication date and the grant date are more linked with the timing of citations (knowledge flow) in the post-reform era by running a simple ordinary least squares (OLS) regression at the individual patent level. Our hypothesis is that the pre-grant publication date, rather than the grant date, is the primary determinant of the earliest citation date for such patent applications. For this estimation, we prepared the following citation data. First, we limited the citation data to those for which the filing date of the cited application is between December 2000 and November 2001. Subsequently, we limited the citations to those for which pre-grant publications were made for the cited applications. In order to eliminate ex-post knowledge, citations that did not satisfy the condition "publication lag <= citation lag" were dropped from our data. Further, we limited the sample to the citing patents that had the earliest effective filing date for each cited patent documents.

The model for the linear regression is as follows: the dependent variable is *Citing_eff_F_T*, which is the effective filing time of the citing patent. There are two main independent variables: *t_pub* is the pre-grant publication time of the cited patent document, and *t-grant* is the granted time of the cited patent document (all measured in calendar month). We used *Di* (technological field dummies, *I* = 1,2,3,4,5) as the control variables.

$$Citing\_eff\_F\_T = \beta_0 + \beta_1 * t\_pub + \beta_2 * t\_grant + Di$$

In order to reduce the effects of truncations, citation lag was limited to the condition citation_lag <= 96. In this model, *t_grant* is likely to be endogenous, given that the applicant may be willing to spend more time for those inventions that need more time for commercialization, while the knowledge flow may occur only late for such applications. However, such endogeneity tends to work against finding support for the hypothesis.

Table 6 presents the summary of the linear regression results. The coefficient of *t_pub* is statistically significant and is close to 1. The coefficient of *t_grant* is almost zero and is less than 1/100 of that of *t_pub*; it is not statistically significant. This result implies that if the publication is delayed by 1 month, the effective filing date of the first citing patent will be delayed by 1 month, while the grant time of the cited patent documents has no influence. One important reason for the finding that the coefficient of *t_pub* is smaller than 1 could be the effect of truncations of the citing patents. We limited the citation lag to citation_lag <= 96; however, the truncation of the citing patents is more serious for the cited patents that are filed later. For those cases, there is a significant possibility that we would not yet observe the potentially citing patents that have the earliest effective filling date but are still under examination and have not been granted a patent.

(Table 6)

(3) Comparing applicant citation flows and examiner citation flows

Figure 8 presents the citation lag dependence of the cumulated citations for applicant citation and examiner citation. The examiner citation curve rises right after the origin (where citation lag = 0), while the applicant citation curve slowly rises after the origin and rises steeply after more than 20 months (where many publications are made). More than 20% of the total examiner citations made within the citation lag of 96 months are made within the 18-month citation lag, while the corresponding ratio of the applicant citations is less than 10%.[8] This difference seems to reflect the difference in the duties of the patent examiners and those of the applicants as well as the difference in the timing when the citations are made. Applicants are primarily obliged to submit what they know at the time of patent application; they are not obliged to search for all the relevant documents disclosed over time that are available at the priority date of the focal patent, while patent examiners have to do so until they make the final decision related to granting the patent. These facts indicate if we use total citations as the measure of knowledge flow, we would get superfluous results about knowledge flow. Thus, it is essential to use applicant citation data to measure knowledge flow.

(Figure 8)

(4) Citation flow volume expansion

Pending patent applications that were eventually abandoned and not granted patents began to be published following the introduction of the pre-grant publication system. We wanted to distinguish self-citations and non-self-citations made for these new publications. However, in many cases, assignee information was lacking for the applications that were not granted and for which only pre-grant publications were made. Figure 9 evaluates the contribution of these new publications to the knowledge flow. The cited patent documents were divided into two groups: (group 1) patent documents of the patent applications that were eventually granted patents (indicated in lavender); (group 2) pre-grant publications for which patents were not granted yet (indicated in

---

[8] Appendix 1 presents the total citations made within the citation lag of 18 months in column A, the total forward citations made within the citation lag of 96 months in column B, and the ratio of A to B in column C for both applicant citations and examiner citations.

green). Figure 9 shows the number of total applicant forward citations made within the 96-month citation lag by a group of cited patent documents.[9] The ratio of citations made by group 2 to total citations varies according to the technological field. It is highest (over 20%) in the field of Computers & Communications and lowest (less than 10%) in the Electrical & Electronic field. Although this number includes self-citations, the third hypothesis was strongly supported.

(Figure 9)

6. Conclusions and Further Research

In this study, we investigated the effects of the introduction of the pre-grant publication system on the diffusion of knowledge. Johnson & Popp (2001) showed that the (total) citation flow begins with publications and suggested that earlier publications would accelerate knowledge flow. However, their study has a fundamental problem of including examiner citations as the measure of knowledge flow, since examiner citations make us observe exactly such correlations, even without any knowledge diffusion effect. In fact, we found that the examiner citation curve begins significantly earlier and more sharply than the applicant citation curve: the shapes of the two curves differ significantly, reflecting the differences in the duties of patent examiners and applicants. This finding supports the need to use applicant non-self-citations as the measure of knowledge flow. In addition, their study used the variation in the grant lag as the source of identification, which could make a causal interpretation difficult. This study used applicant (inventor) non-self-citations (excluding examiner citations) to track knowledge flow and identified the causal effect of disclosure by examining the changes in behaviors before and after the legal reform.

Further, this study examined the effects of the pre-grant publication system on the pattern of knowledge diffusion in order to explore the relative importance of the two potential effects of disclosure—avoiding duplications in R&D and/or patenting and accelerating the increase of knowledge stock for further inventions—. If duplicative inventions are important, early disclosure of the focal patent increases the initial slope of citation since it makes an inventor with a duplicative invention recognize the patent applications more at the time of its patent application. In addition, if it helps reducing duplicative R&D and/or patenting in the long run, the initial increase will be associated with the decline of the long-run level of citation flows from the focal patent. In contrast, if early discloser promotes the rapid expansion of knowledge stock for complementary

[9] The citation lag dependence of the number of total applicant citations by the group of cited patent documents is shown in Appendix 2.

inventions, it increases both the initial slope as well as the long-term level of citation flow. The duplication reduction effect of early disclosure is expected to be stronger in the fields where the disclosure was late before the legal change. This issue has not been explored in the extant literature.

We tested three hypotheses in this study:

(1) The introduction of the pre-grant publication system accelerated the knowledge diffusion, especially in technological fields where the grant lag was long.

(2) The effects of pre-grant publications on the diffusion curve can be heterogeneous, depending on the relative importance of the two citation flows: those for duplicative inventions and those for complementary inventions.

(3) The pre-grant publication system increases the sources of knowledge flows since the pending patent applications that were eventually abandoned and not granted patents were also published under this system.

The results of the nonlinear regression analysis show that knowledge diffusion began significantly earlier following the introduction of the pre-grant publication system. Except in the Chemical field, there is highly significant evidence for the effects of pre-grant publication on the acceleration of knowledge diffusion (five out of six fields). The effect was the strongest in the Computers & Communications field (where the publication lag was the longest before the reform), and it was the second strongest in the Drugs and Medical field.

In addition, the initial slope of the diffusion curve rose while the long-term level of citation flow declined in the Computers & Communications field, which had the longest publication lag before the reform. In contrast, both of them rose in the Electrical & Electronic field. These results suggest the possibility that early disclosure not only stimulated complementary inventions but also helped inventors recognize early the duplications and then helped the reductions of either duplicative R&D themselves and/or of duplicative applications, in a field with a long publication lag.

Additionally, the publications of the pending patent applications that were eventually abandoned and not granted patents significantly contributed to the expansion of the knowledge flow.

Since AIPA reform introduced a pre-grant publication system that allows the applicants of purely domestic applications to forgo pre-grant publication, it did not constrain the applicants who opt for secrecy before a grant. As to the published applications before the grant, our study convincingly show that the knowledge diffusion from them accelerated after the legal change, especially in those fields where the grant lag was large. Furthermore, there are significant citation flows to the non-granted patent applications. Thus, the introduction of the pre-grant publication under the AIPA reform seems to have had positive effects on innovation.

There are a number of other issues that could be examined in future research. Since the applicant's duty to candor continues until the patent is issued, strictly speaking, the IDS should contain not only the applicant's knowledge before filing but also what he/she gets to know after filing the patent application. All the ex-post knowledge has not been eliminated from our applicant citation data. It would be interesting to identify exactly how ex-post knowledge is included in the applicant citation data by examining the IDS documents. Another aspect that requires a separate investigation is the effect of the introduction of the pre-grant publication system on internal knowledge sharing (measured by self-citation) in a firm. If the applicant is an individual person or a small enterprise, the knowledge flow within the entity would not be influenced much by the introduction of the pre-grant publication system. However, in the case of a large company, the pre-grant publication system might promote the knowledge diffusion within the organization. Finally, the pre-grant publication system may have helped the examination process of the patent offices by facilitating the sharing of more accurate information on the patent search results across examiners. Future research could explore these additional knowledge flow effects of the pre-grant publication system.

Acknowledgment

This project was supported by the JSPS Grant-in-Aid for Scientific Research JSPS(B:26285055). We would like to thank JSPS for its granting aid. We would like to thank Professor Mark Schankerman, Professor Stuart Graham and Professor Tetsuo Wada for giving comments to the earlier versions of this paper. We also would like to thank research assistant Junko Morikawa for her help in file wrapper investigation.

References

Alcacer, J., & Gittelman, M. (2004). How do I know what you know? Patent examiners and the generation of patent citations. Available at SSRN: http://ssrn.com/abstract=548003 or http://dx.doi.org/10.2139/ssrn.548003.

Alcacer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. The Review of Economics and Statistics, 88(4), 774-779.

Alcacer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in US patents: An overview and analysis. Research Policy, 38(2), 415-427.

Allison, J. R., & Lemley, M. A. (1998). Empirical evidence on the validity of litigated patents. American Intellectual Property Law Association (AIPLA) Quarterly Journal, 26, 185-, SSRN: http://ssrn.com/abstract=118149.

Buchanan, J. M. (2006). Deference overcome: Courts' invalidation of patent claims as anticipated by art considered by the PTO. 2006 STAN. TECH. L. REV. 2 https://journals.law.stanford.edu/sites/default/files/stanford-technology-law-review/online/buchananj-deference.pdf

Caballero, R. J., & Jaffe, A. B. (1993). How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In O. J. Blanchard and S. Fischer, eds., NBER Macroeconomics Annual, pp. 15-86. Cambridge, Mass.: MIT Press.

Carley, M., Hegde, D., & Marco, A. (2013). What is the probability of receiving a US patent? USPTO Economic Working Paper No. 2013-2. http://www.uspto.gov/ip/officechiefecon/OCE_WP_2013-2.pdf.

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

Gans, J. S., Hsu, D. H., & Stern, S. (2008). The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. Management Science, 54(5), 982-997.

Graham, S. J. H. & Hegde, D. (2012). Do inventors value secrecy in patenting? Evidence from the American Inventor's Protection Act of 1999. Available at SSRN: http://ssrn.com/abstract=2170555 or http://dx.doi.org/10.2139/ssrn.2170555.

Graham, S., & Hegde, D. (2015). Disclosing patents' secrets. Science 347(6219), 236-237.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). National Bureau of Economic Research.

Hegde, D., & Luo, H. (2013). Imperfect information, patent publication, and the market for ideas. Harvard Business School Strategy Unit Working Paper (14-019).

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics, 108(3), 577-598.

Jaffe, A. B., Fogarty, M. S., & Banks, B. A. (1998). Evidence from patent citations on the impact of NASA and other federal labs on commercial innovation. Journal of Industrial Economics, 46, 183-206.

Johnson, D. K. N., & Popp, D. (2001). Forced out of the closet: The impact of the American Inventors Protection Act on the timing of patent disclosure (No. w8374). National Bureau of Economic Research.

Juneau, T. L., & MacAlpine, J. K. (2000). Protecting patents from the beginning: The importance of information disclosure statements during patent prosecution. Journal of the Patent and Trademark Office Society, 82(8), 577-595.

KSR International, Co. v. Teleflex, Inc. (2007). 127 S. Ct. 1727 (2007) (No. 04-1350), 22-23.

Figure 1 Change in timing of publication before and after the AIPA reform

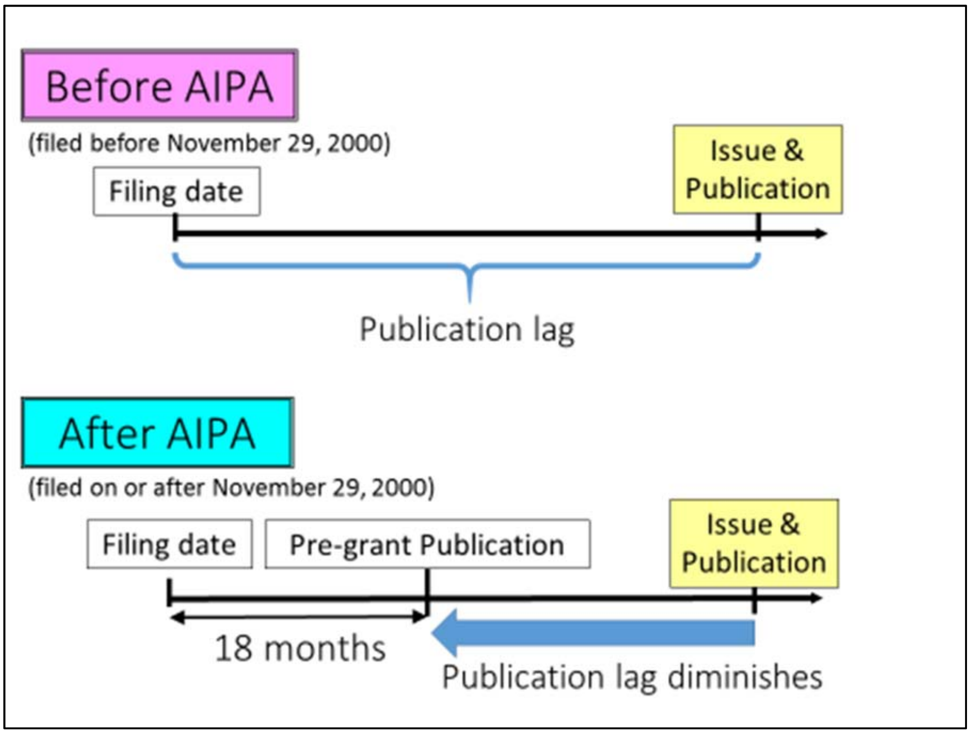


Figure 2 Cases where patent applications are filed in foreign counties for the same invention

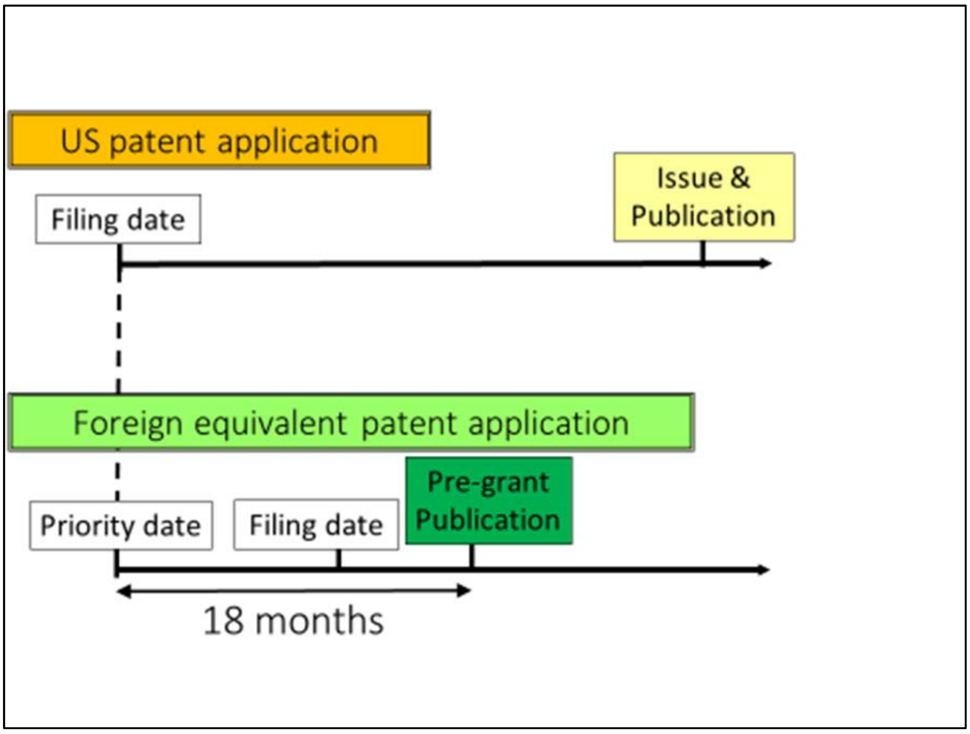


Figure 3 Change in application coverage before and after the AIPA reform

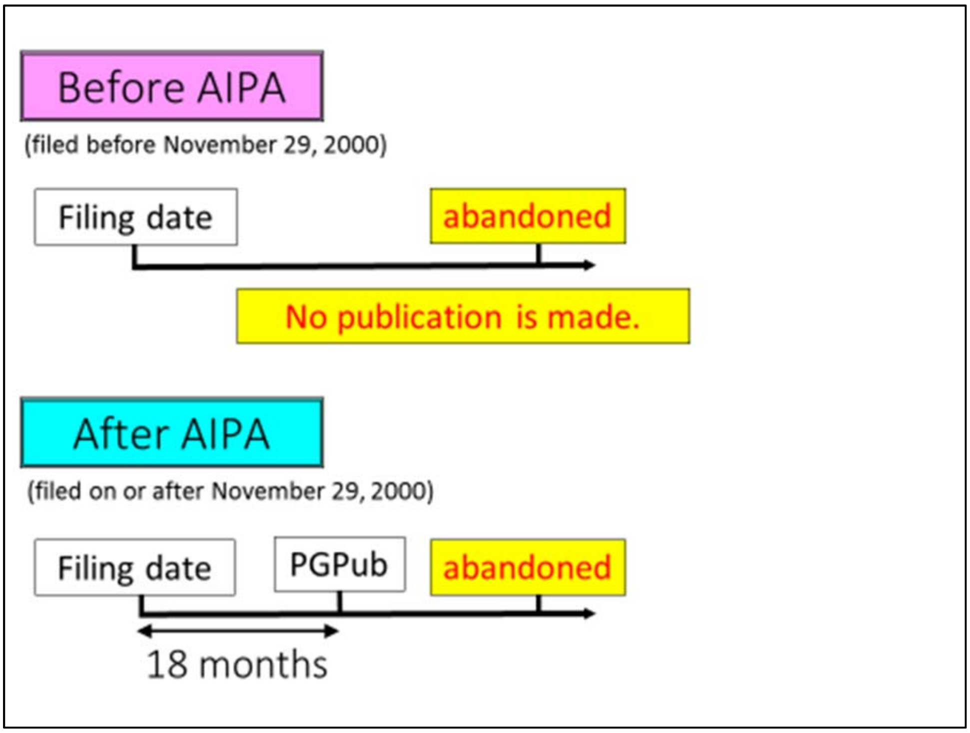
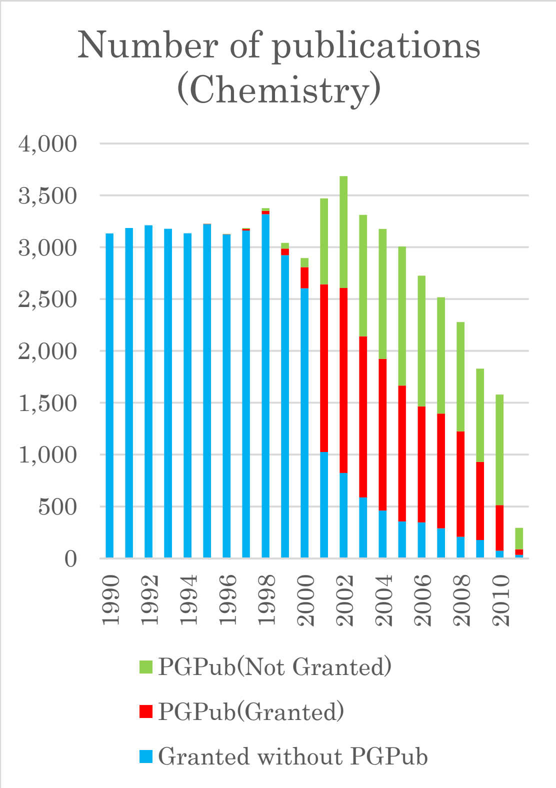
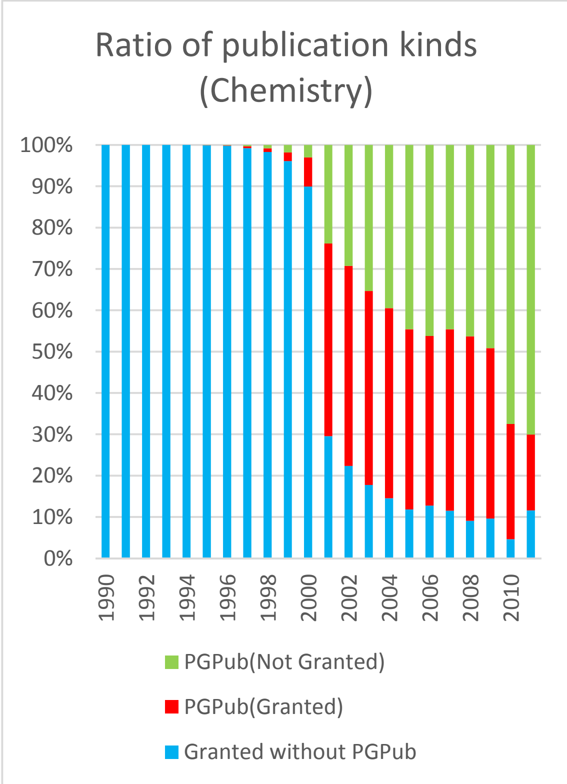


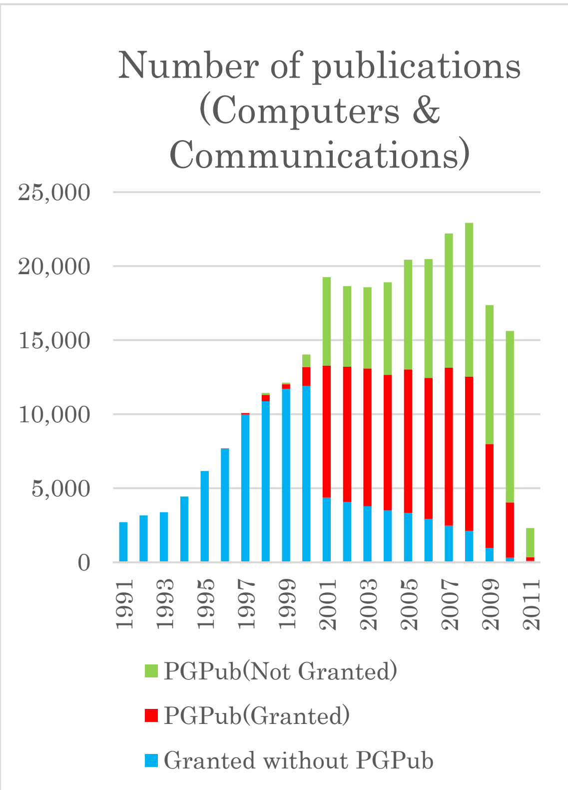
Figure 4 Number and composition of publications of purely domestic normal applications over the filing years from 1990 to 2011

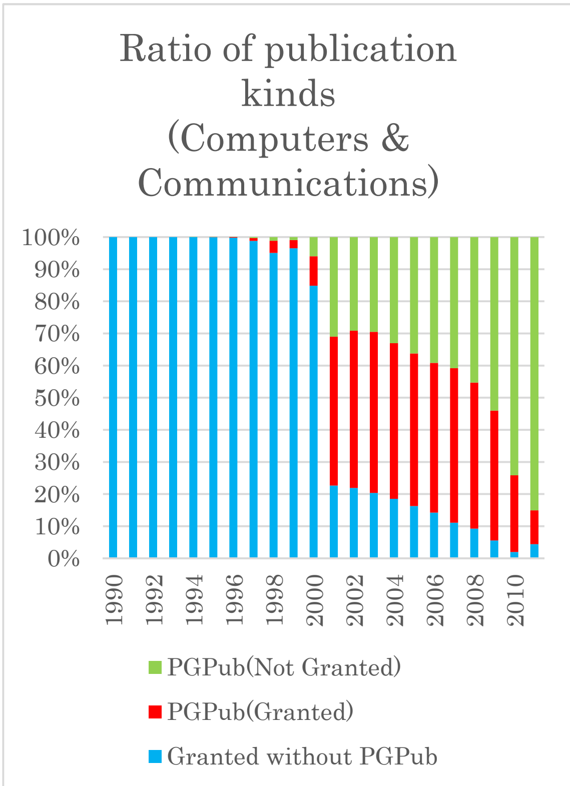
A



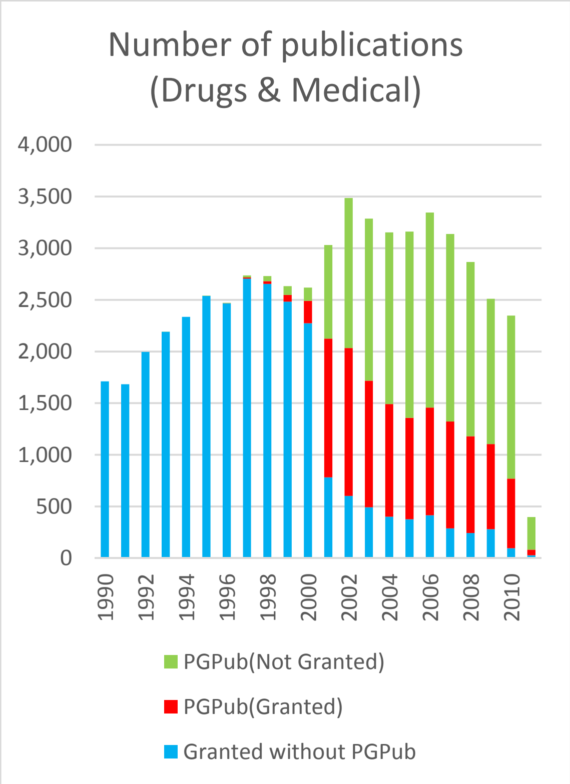


B

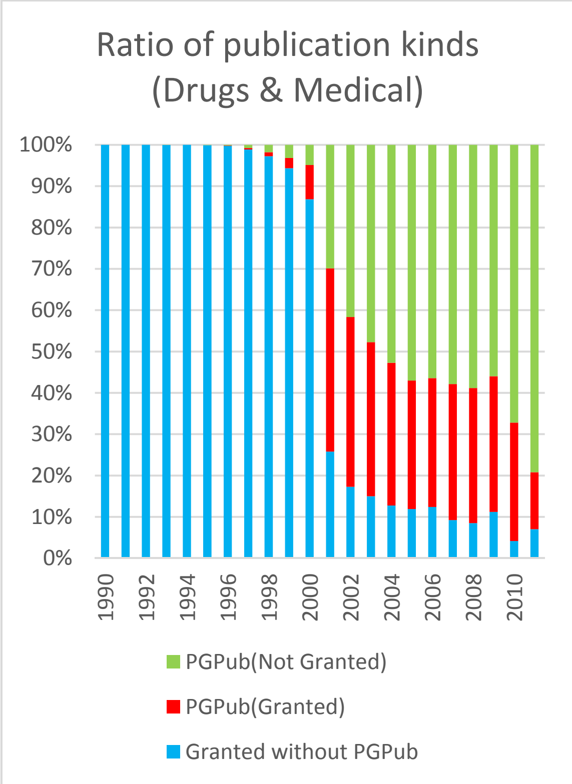




C

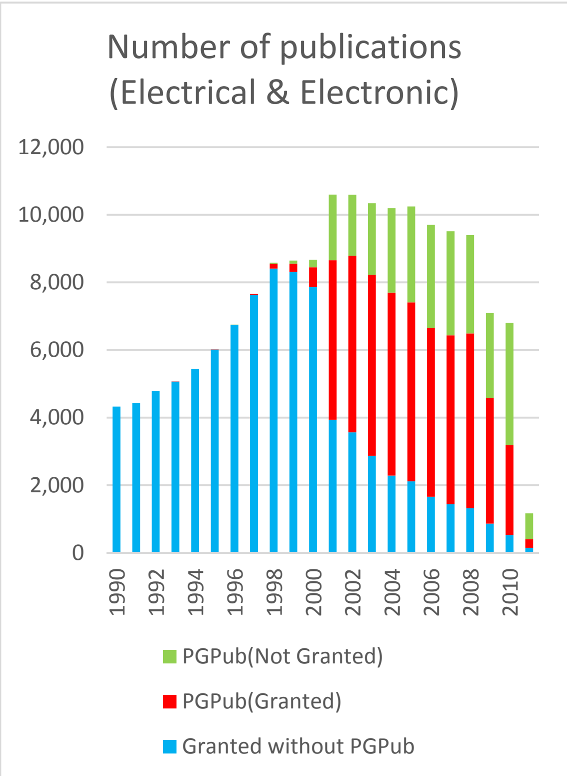


Number of publications
(Drugs & Medical)
4,000
3,500
3,000
2,500
2,000
1,500
1,000
500
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub

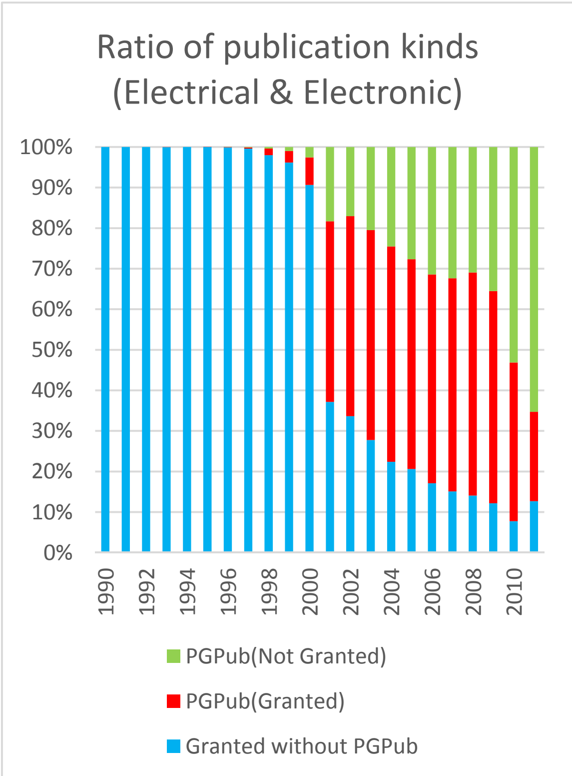


Ratio of publication kinds
(Drugs & Medical)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub

D

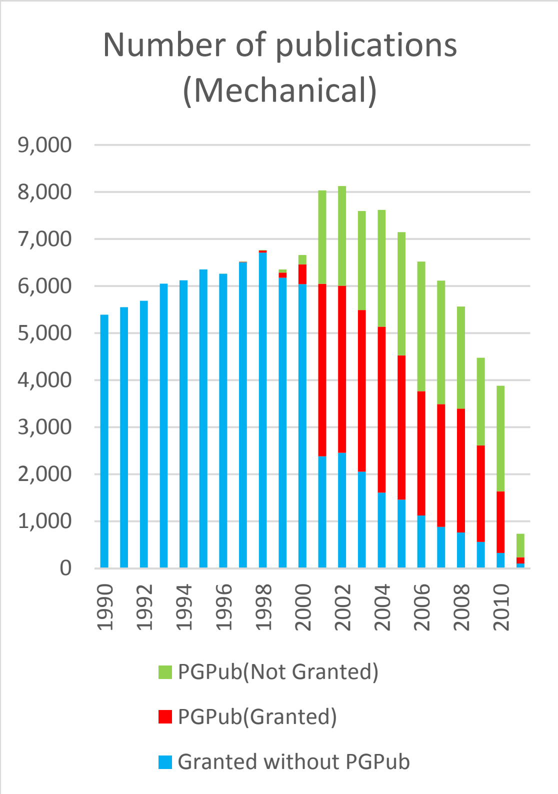


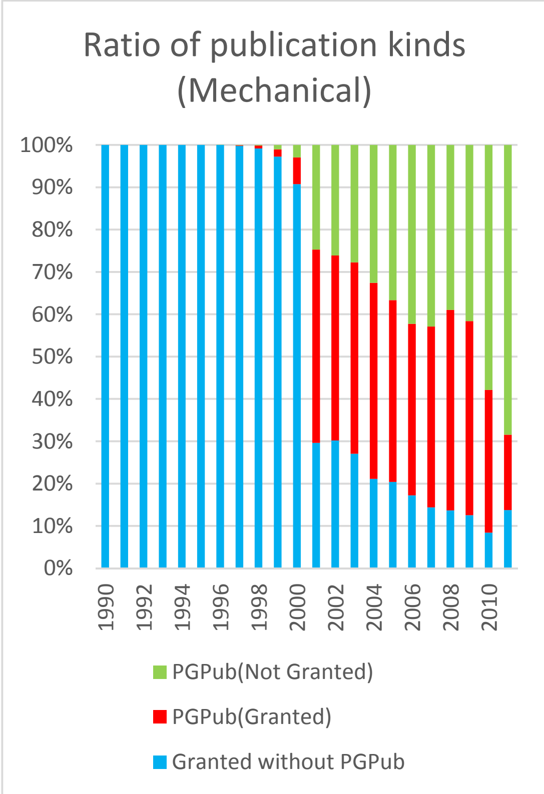
Number of publications
(Electrical & Electronic)
12,000
10,000
8,000
6,000
4,000
2,000
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub



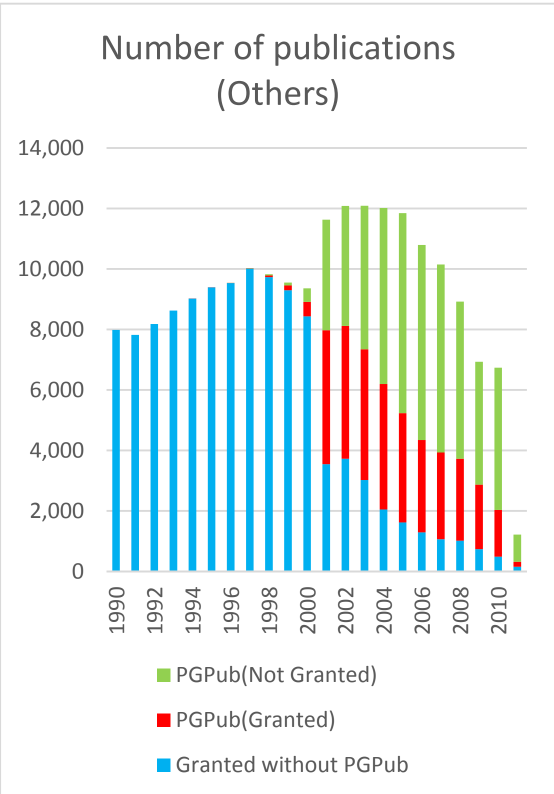
Ratio of publication kinds
(Electrical & Electronic)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub

E


Number of publications
(Mechanical)
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub


Ratio of publication kinds
(Mechanical)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub

F


Number of publications
(Others)
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub

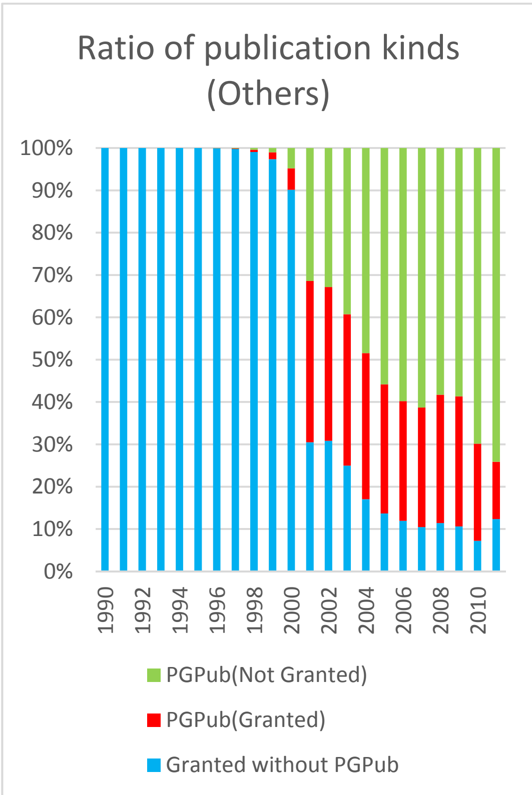

Ratio of publication kinds
(Others)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
PGPub(Not Granted)
PGPub(Granted)
Granted without PGPub

Figure 5 Citation origin and overlap of U.S. patent references

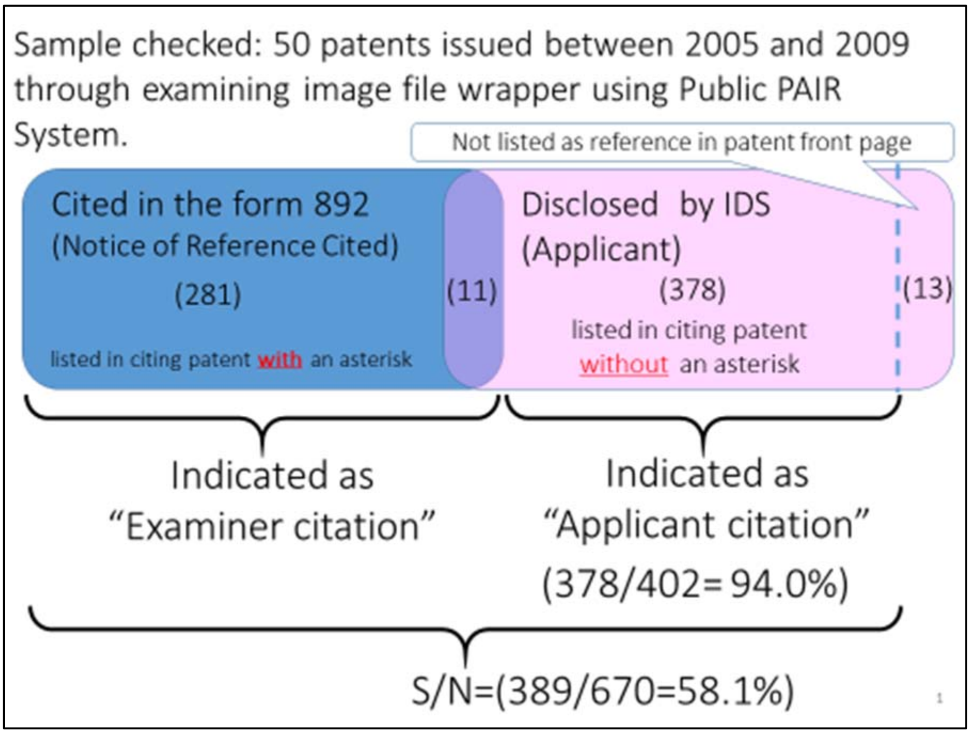
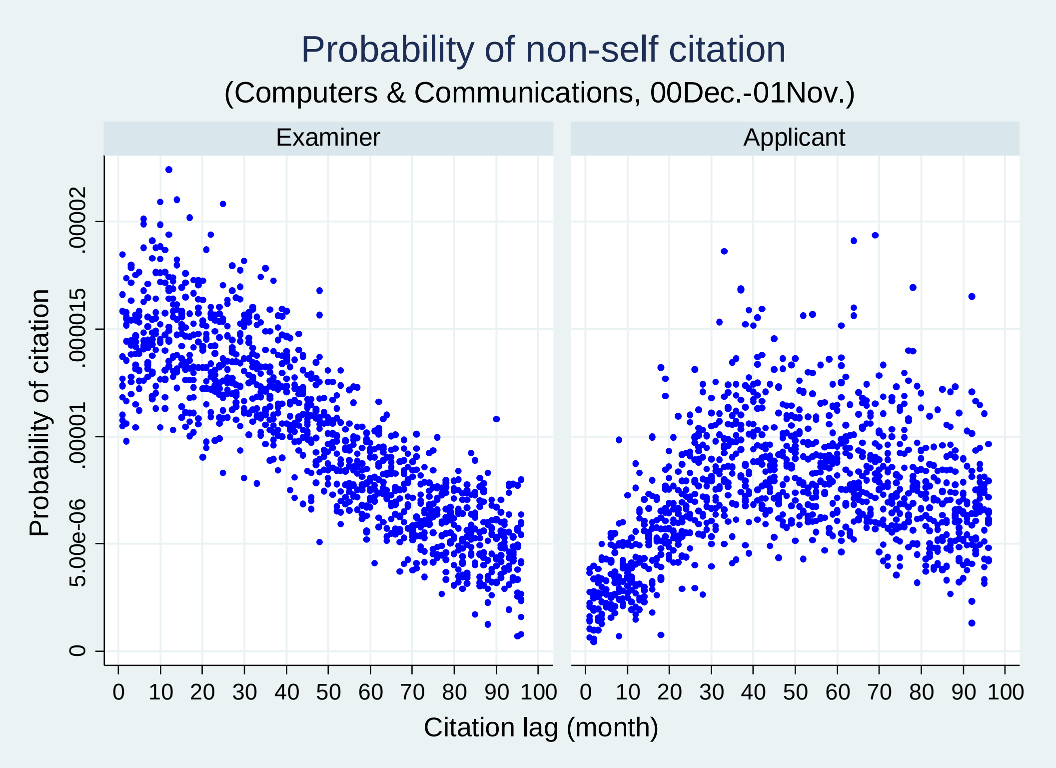


Figure 6 Probability of applicant citation and examiner citation

A Computers & Communications



B Electrical & Electronic

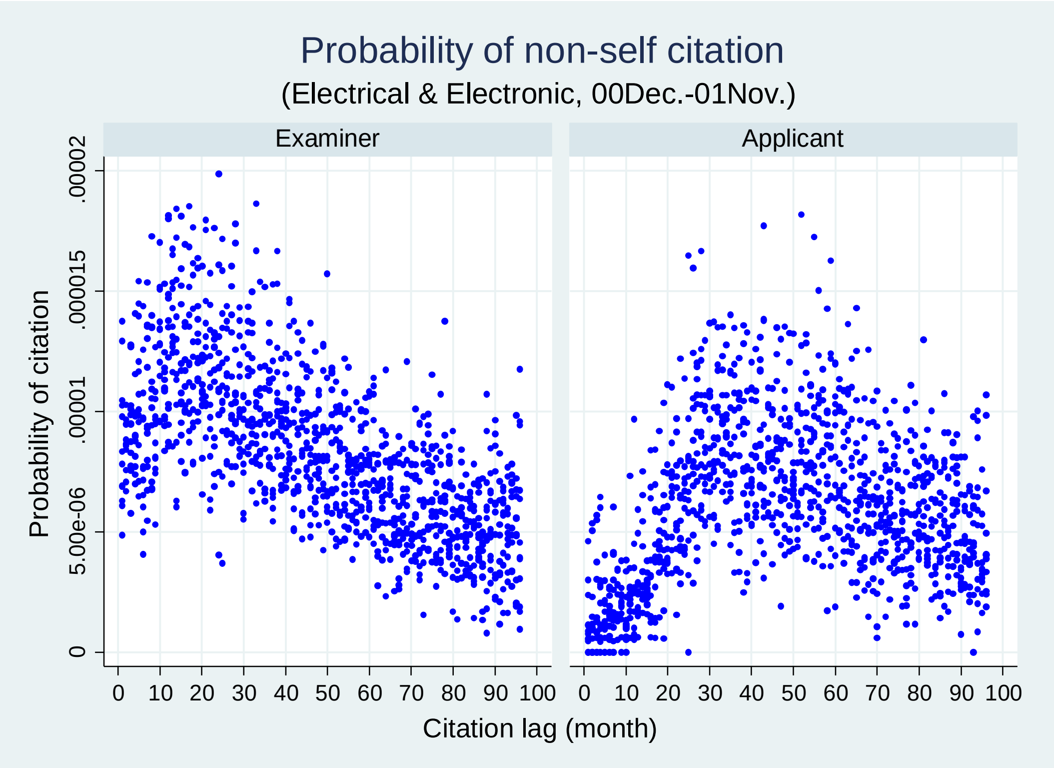
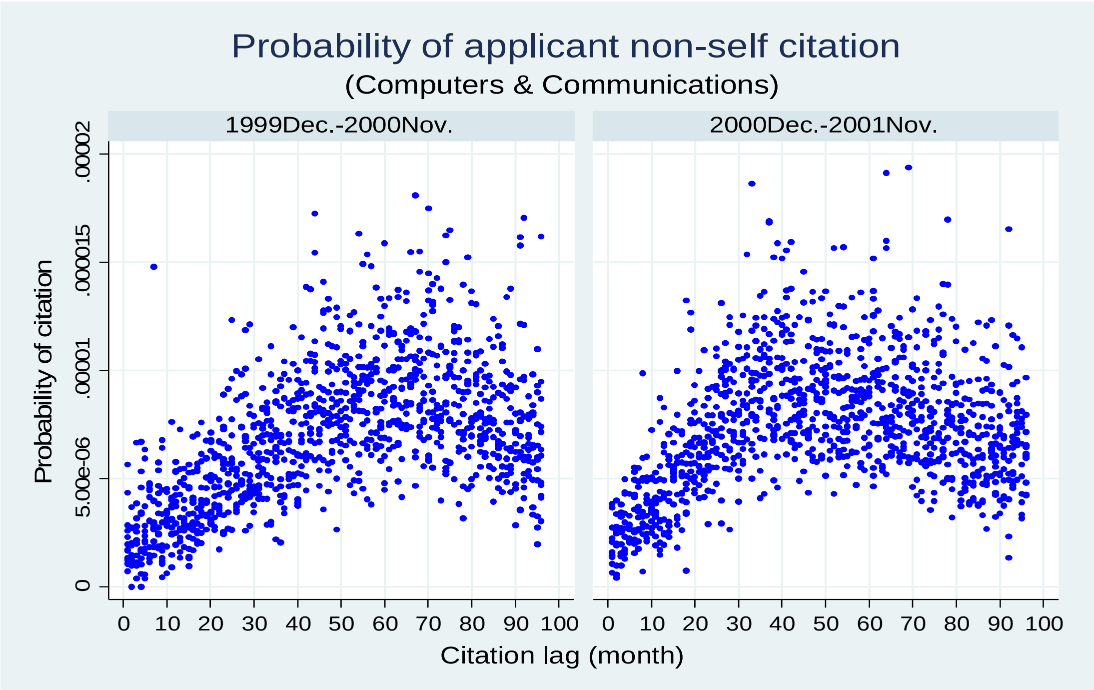


Figure 7 Probability of non-self-citation in Computers & Communications before and after law revision

A　Actual data only



B　Actual data plotted with estimated value (red) by regression in model 3

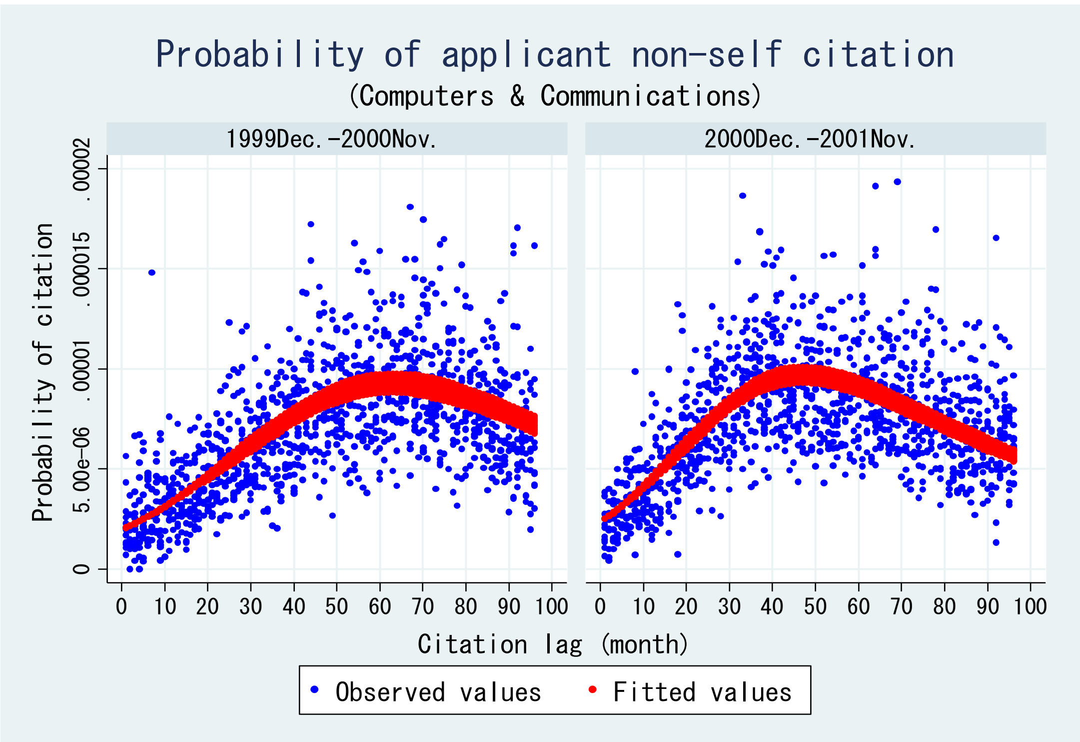
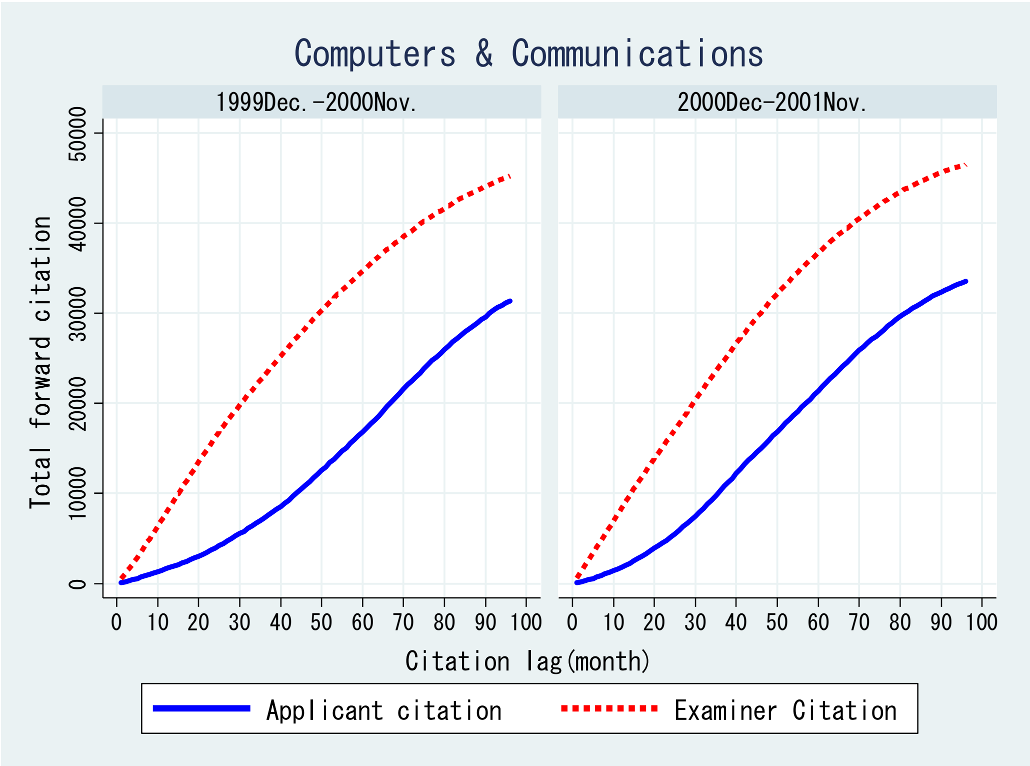


Figure 8 Total number of applicant citations and examiner citations vs. citation lag

A



B

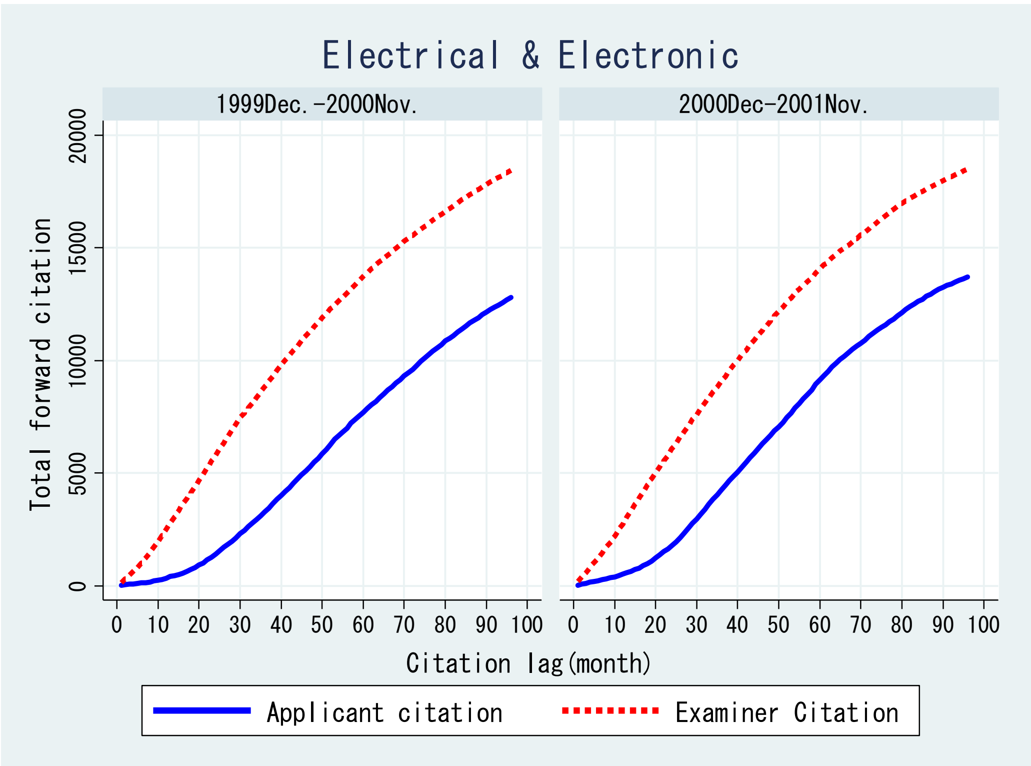


Figure 9 Total applicant forward citations made in 96-month citation lag by group of cited patent documents

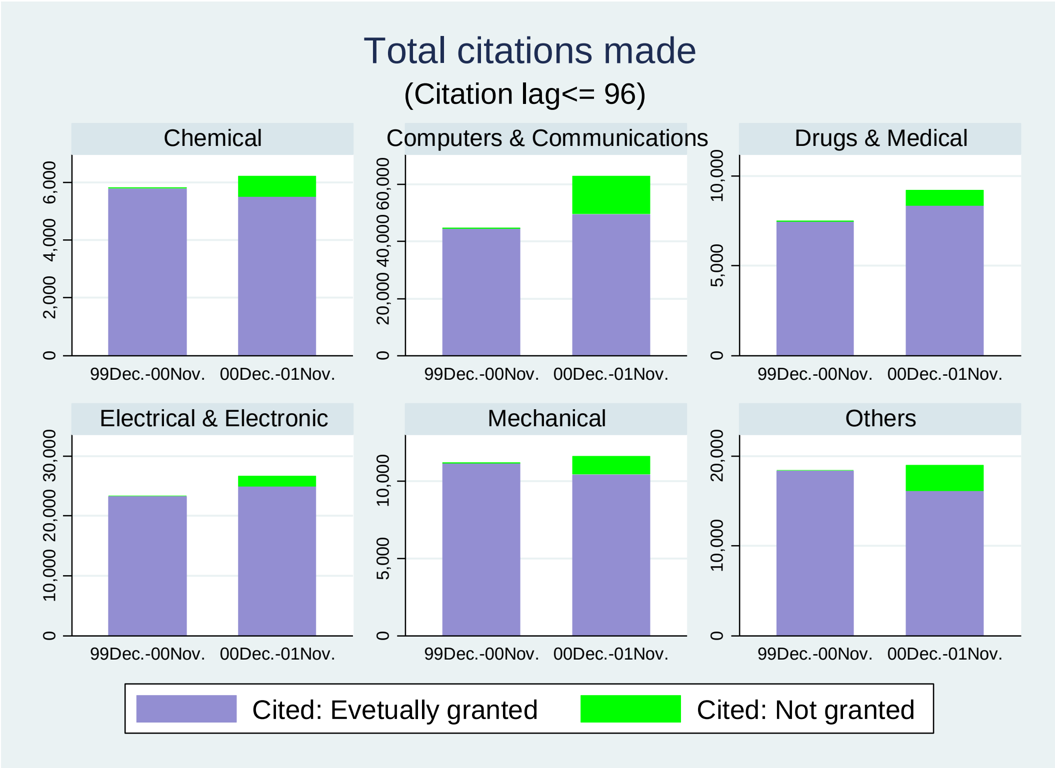


Table 1 Statistics for granted purely domestic applications claiming no priority based on earlier applications

term	Statistics	Chemical	Computers & Communications	Drugs & Medical	Electrical & Electronic	Mechanical	Others
97Dec. – 98Nov.	A: mean Publication lag	24.72	36.03	25.04	27.27	23.89	22.93
	B: mean Grant lag	24.80	36.53	25.12	27.44	23.97	23.01
	B – A	0.08	0.50	0.08	0.17	0.09	0.08
	sd pub lag	9.80	15.01	11.94	11.59	9.04	9.86
	n	3,298	11,124	2,670	8,662	6,738	9,802
98Dec. – 99Nov.	A: mean Publication lag	24.53	42.32	25.48	26.29	22.88	23.08
	B: mean Grant lag	24.78	42.76	25.81	26.61	23.03	23.26
	B – A	0.25	0.45	0.33	0.32	0.16	0.18
	sd pub lag	10.55	18.32	12.61	12.14	9.59	11.26
	n	3,059	11,990	2,542	8,627	6,356	9,485
99Dec. – 00Nov.	A: mean Publication lag	26.12	48.54	29.20	26.84	24.23	25.03
	B: mean Grant lag	26.34	49.01	29.50	27.05	24.38	25.15
	B – A	0.22	0.47	0.30	0.20	0.15	0.12
	sd pub lag	12.00	22.03	14.09	12.49	11.02	12.55
	n	2,823	12,981	2,433	8,491	6,515	9,121
00Dec. – 01Nov.	A: mean Publication lag	19.75	28.76	21.64	19.75	19.28	19.49
	B: mean Grant lag	27.37	50.97	31.08	26.33	25.54	26.36
	B – A	7.63	22.21	9.44	6.58	6.26	6.87
	sd pub lag	7.40	20.33	9.78	7.96	7.36	8.99
	n	2,684	13,608	2,246	8,552	6,069	7,969
00Dec. – 01Nov.	A: mean Publication lag	20.29	28.14	22.04	19.88	19.61	20.03
	B: mean Grant lag	30.90	50.45	37.33	27.50	26.75	27.94
	B – A	10.61	22.31	15.29	7.62	7.14	7.91
	sd pub lag	9.04	19.98	11.59	8.09	7.70	9.53
	n	2,533	13,167	1,998	8,830	5,978	8,012
02Dec. – 03Nov.	A: mean Publication lag	20.65	27.63	22.85	20.22	19.93	20.47
	B: mean Grant lag	33.80	51.17	43.10	29.42	28.78	30.79
	B – A	13.15	23.54	20.25	9.20	8.85	10.32
	sd pub lag	9.43	19.39	12.50	8.42	8.18	9.82
	n	2,217	12,826	1,715	8,191	5,525	7,446

Table 2 Statistics for purely domestic normal applications by NBER-identified assignees that were granted patents

term	Statistics	Chemical	Computers & Communications	Drugs & Medical	Electrical & Electronic	Mechanical	Others
97Dec. – 98Nov.	A: mean Publication lag	25.31	36.24	25.56	27.77	24.97	25.12
	B: mean Grant lag	25.40	36.74	25.68	27.94	25.05	25.22
	B – A	0.08	0.50	0.12	0.18	0.07	0.10
	sd pub lag	9.96	14.42	11.71	11.56	9.48	10.70
	n	2,502	10,349	1,766	7,511	4,134	4,460
98Dec. – 99Nov.	A: mean Publication lag	25.06	42.57	26.26	26.68	24.01	25.95
	B: mean Grant lag	25.34	43.02	26.63	27.03	24.23	26.17
	B – A	0.28	0.44	0.37	0.35	0.22	0.21
	sd pub lag	10.33	17.43	12.43	12.14	10.09	12.61
	n	2,351	11,118	1,662	7,514	3,980	4,452
99Dec. – 00Nov.	A: mean Publication lag	26.89	47.96	30.35	27.17	25.64	27.76
	B: mean Grant lag	27.10	48.39	30.69	27.39	25.82	27.92
	B – A	0.21	0.43	0.34	0.22	0.18	0.16
	sd pub lag	11.43	20.41	14.12	12.38	11.31	13.11
	n	2,169	11,838	1,615	7,363	4,123	4,492
00Dec. – 01Nov.	A: mean Publication lag	20.11	28.27	22.31	19.95	19.82	20.43
	B: mean Grant lag	27.84	49.77	32.26	26.60	27.01	29.23
	B – A	7.74	21.51	9.95	6.65	7.18	8.80
	sd pub lag	7.45	19.18	9.76	8.13	7.67	9.69
	n	2,108	12,423	1,572	7,525	3,861	3,994
00Dec. – 01Nov.	A: mean Publication lag	20.22	27.42	22.01	19.91	19.74	20.59
	B: mean Grant lag	31.24	48.57	37.81	27.54	27.85	30.82
	B – A	11.02	21.14	15.80	7.63	8.11	10.23
	sd pub lag	8.68	18.68	11.14	7.93	7.51	10.03
	n	2,002	11,657	1,295	7,750	3,701	4,015
02Dec. – 03Nov.	A: mean Publication lag	20.45	26.74	22.86	20.15	19.59	20.44
	B: mean Grant lag	33.74	49.10	42.61	29.31	29.29	32.57
	B – A	13.29	22.36	19.75	9.16	9.70	12.13
	sd pub lag	8.86	18.16	11.68	7.90	7.19	9.71
	n	1,641	10,915	1,053	7,116	3,424	3,615

Table 3 Descriptive statistics of data used for regression analysis

Probability of nonself applicant citation

1999Dec.-2000Nov.

Technological fields	N	mean	sd	p50	min	max	Number of P=0 points	Ratio of P=0 points
Chemical	1152	8.69E-06	8.76E-06	6.59E-06	0	0.000066	307	26.6%
Computers & Communications	1152	6.95E-06	3.28E-06	6.80E-06	0.00E+00	2.24E-05	2	0.2%
Drugs & Medical	1152	2.76E-05	2.43E-05	2.21E-05	0	0.000157	170	14.8%
Electrical & Electronic	1152	5.80E-06	3.31E-06	5.59E-06	0	2.43E-05	28	2.4%
Mechanical	1152	5.73E-06	4.67E-06	5.01E-06	0	0.000038	144	12.5%
Others	1152	7.56E-06	5.54E-06	7.09E-06	0	5.05E-05	98	8.5%
Total	6912	1.04E-05	1.35E-05	6.87E-06	0	0.000157	749	10.8%

2000Dec.-2001Nov.

Technological fields	N	mean	sd	p50	min	max	Number of P=0 points	Ratio of P=0 points
Chemical	1152	8.74E-06	9.34E-06	6.36E-06	0	6.96E-05	342	29.7%
Computers & Communications	1152	7.34E-06	3.01E-06	7.16E-06	4.15E-07	1.94E-05	0	0.0%
Drugs & Medical	1152	3.39E-05	3.94E-05	2.68E-05	0	0.00089	149	12.9%
Electrical & Electronic	1152	6.12E-06	3.34E-06	5.99E-06	0	2.13E-05	17	1.5%
Mechanical	1152	6.27E-06	5.30E-06	5.15E-06	0	3.95E-05	150	13.0%
Others	1152	8.07E-06	6.01E-06	7.14E-06	0	3.54E-05	101	8.8%
Total	6912	1.17E-05	1.96E-05	7.22E-06	0	0.00089	759	11.0%

sigma N

1999Dec.-2000Nov.

Technological fields	N	mean	sd	p50	min	max
Chemical	1152	100298	61030	102717	1197	204640
Computers & Communications	1152	246752	170711	237702	1826	582321
Drugs & Medical	1152	91810	62074	89760	807	208830
Electrical & Electronic	1152	181263	115035	178052	1785	398067
Mechanical	1152	130872	79159	132440	1493	268482
Others	1152	138660	87443	137605	1621	299047
Total	6912	148276	115825	131469	807	582321

2000Dec.-2001Nov.

Technological fields	N	mean	sd	p50	min	max
Chemical	1152	107642	60572	111237	1432	209200
Computers & Communications	1152	290917	181206	286460	2237	633079
Drugs & Medical	1152	105359	64946	105436	1016	225474
Electrical & Electronic	1152	201268	119083	198039	2352	421056
Mechanical	1152	140122	79000	141879	1895	274522
Others	1152	151854	89841	151465	1830	312712
Total	6912	166194	125147	146444	1016	633079

Table 4A Summary of nonlinear regression results

Coefficients of Independent Variables (no intercept)	Dependent Variable: P,co,i,t,i,T								
	Computers & comuunications			Electrical & Electronic			Drugs & Medical		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
$\alpha 1$	.0000256***	.0000305***	.0000321***	.0000135***	.0000138***	.0000128***	.0000436***	.0000481***	.0000420***
	(2.82e-06)	(4.23e-06)	(5.28e-06)	(6.90e-07)	(6.93e-07)	(6.11e-07)	(7.37e-06)	(7.21e-06)	(6.42e-06)
$\alpha 2$	-3.81e-06***		6.13e-06***	8.09e-07**		1.59e-06***	7.27e-06***		.0000105***
	(1.05e-06)		(1.60e-06)	(3.38e-07)		(2.76e-07)	(2.26e-06)		(2.05e-06)
$\mu 1$	.0569***	.0612***	.0571***	.123***	.135***	.138***	.157***	.157***	.151***
	(.00217)	(.00174)	(.00145)	(.00706)	(.00642)	(.0067)	(.0271)	(.0198)	(.0185)
$\mu 2$	.0194***			.0253**			.0114		
	(.00385)			(.0106)			(.0311)		
$\tau 1$	43.7***	45.4***	44.8***	24.0***	24.4***	21.8***	26.2***	27.9***	22.7***
	(2.74)	(2.91)	(3.63)	(.942)	(.830)	(.737)	(2.21)	(1.80)	(1.84)
$\tau 2$	-16.0***	-8.87***		-3.34***	-4.25***		-6.05***	-8.78***	
	(1.62)	(.505)		(.827)	(.585)		(1.58)	(1.06)	
λ	2.20e-06***	2.68e-06***	2.89e-06***	2.68e-06***	2.65e-06***	2.60e-06***	1.27e-06	1.37e-06	1.26e-06
	(1.95e-07)	(2.54e-07)	(3.01e-07)	(1.72e-07)	(1.74e-07)	(1.70e-07)	(1.19e-06)	(1.19e-06)	(1.17e-06)
Observations	2,304	2,304	2,304	2,304	2,304	2,304	2,304	2,304	2,304
Log Likelihood	26665	26652	26589	26433	26422	26424	20688	20681	20682

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.10

Table 4B Summary of nonlinear regression results

Coefficients of Independent Variables (no intercept)	Dependent Variable: P, co, i, t, i, T								
	Mechanical			Chemical			Others		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
$\alpha 1$	.0000112***	.0000112***	.0000103***	.0000135***	.0000130***	.0000128***	.0000149***	.0000150***	.0000144***
	(1.02e-06)	(9.92e-07)	(8.35e-07)	(1.81e-06)	(1.67e-06)	(1.59e-06)	(1.29e-06)	(1.29e-06)	(1.24e-06)
$\alpha 2$	6.69e-07		1.23e-06***	-7.46e-07		-2.08e-07	2.39e-07		1.29e-06***
	(4.59e-07)		(3.86e-07)	(6.76e-07)		(5.01e-07)	(5.41e-07)		(4.46e-07)
$\mu 1$	.126***	.150***	.158***	.126***	.136***	.138***	.116***	.119***	.117***
	(.0117)	(.0119)	(.0126)	(.0158)	(.0142)	(.0146)	(.00857)	(.00743)	(.00745)
$\mu 2$	.046**			.0165			.00690		
	(.0202)			(.0231)			(.0123)		
$\tau 1$	24.8***	24.7***	22.4***	25.5***	24.7***	23.6***	28.4***	28.5***	26.1***
	(1.46)	(1.20)	(1.08)	(2.10)	(1.89)	(1.73)	(1.44)	(1.37)	(1.37)
$\tau 2$	-2.92**	-3.56***		-3.23*	-1.95		-4.80***	-5.07***	
	(1.26)	(.888)		(1.71)	(1.23)		(1.23)	(.806)	
λ	2.43e-06***	2.29e-06***	2.15e-06***	8.82e-07	8.18e-07	6.81e-07	1.91e-06***	1.90e-06***	1.90e-06***
	(4.70e-07)	(4.64e-07)	(4.48e-07)	(8.67e-07)	(8.51e-07)	(8.41e-07)	(3.90e-07)	(3.89e-07)	(4.00e-07)
Observations	2,304	2,304	2,304	2,304	2,304	2,304	2,304	2,304	2,304
Log Likelihood	25253	25247	25250	23862	23861	23860	25082	25081	25073

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 5 Number of citations made before publication of cited applications vs. number of citations made on and after publication of cited applications

Technological fields	Filling year month	A: Publication lag > citation lag	B: Publication lag <= citation lag	share of A
Chemical	1999Dec.-2000Nov.	176	2045	7.9%
Chemical	2000Dec-2001Nov.	122	1802	6.3%
Computers & Communications	1999Dec.-2000Nov.	9,450	21923	30.1%
Computers & Communications	2000Dec-2001Nov.	5,296	28242	15.8%
Drugs & Medical	1999Dec.-2000Nov.	399	3591	10.0%
Drugs & Medical	2000Dec-2001Nov.	343	4139	7.7%
Electrical & Electronic	1999Dec.-2000Nov.	1,124	11658	8.8%
Electrical & Electronic	2000Dec-2001Nov.	888	12812	6.5%
Mechanical	1999Dec.-2000Nov.	298	3870	7.1%
Mechanical	2000Dec-2001Nov.	199	3770	5.0%
Others	1999Dec.-2000Nov.	500	4715	9.6%
Others	2000Dec-2001Nov.	354	4189	7.8%

Table 6 Summary of linear regression results

Source	SS	df	MS			
Model	132599.127	7	18942.7324	Number of obs.	=	5396
				F(7, 5388)	=	58.97
				Prob > F	=	0.0000
Residual	1730878.97	5388	321.247025	R-squared	=	0.0712
				Adj R-squared	=	0.0700
Total	1863478.1	5395	345.408359	Root MSE	=	17.923

Citing_eff_F_T	Coef.	Std. Err.	t	P > t	[95% Conf.	Interval]
t_pub: Cited First Publication time(month)	0.883	0.0539	16.36	0.000	0.777	0.988
t_grant: Cited Grant time(month)	-0.00747	0.0151	-0.50	0.621	-0.0371	0.0221
Technological field dummy						
Computers & Communications	-7.70	1.09	-7.07	0.000	-9.83	-5.56
Drugs & Medical	-5.89	1.31	-4.51	0.000	-8.45	-3.33
Electrical & Electronic	-4.12	1.11	-3.71	0.000	-6.29	-1.94
Mechanical	-1.46	1.22	-1.20	0.231	-3.86	0.933
Others	-0.486	1.20	-0.41	0.685	-2.83	1.86
_cons	28.60	1.62	17.67	0.000	25.43	31.78

Appendix 1 Applicant citations and examiner citations made in 18-month citation lag and 96-month citation lag

Technological Field	Filing date	citer	A: within 18 months	B: within 96 months	Ratio of A/B
Computers & Communications	1999Dec.-2000Nov.	Applicant	2633	31373	8.4%
		Examiner	12089	45216	26.7%
	2000Dec-2001Nov.	Applicant	3276	33538	9.8%
		Examiner	12487	46476	26.9%
Electrical & Electronic	1999Dec.-2000Nov.	Applicant	723	12782	5.7%
		Examiner	4128	18425	22.4%
	2000Dec-2001Nov.	Applicant	988	13700	7.2%
		Examiner	4455	18490	24.1%
Drugs & Medical	1999Dec.-2000Nov.	Applicant	171	3990	4.3%
		Examiner	420	2015	20.8%
	2000Dec-2001Nov.	Applicant	290	4482	6.5%
		Examiner	483	2116	22.8%
Mechanical	1999Dec.-2000Nov.	Applicant	211	4168	5.1%
		Examiner	1538	6612	23.3%
	2000Dec-2001Nov.	Applicant	218	3969	5.5%
		Examiner	1372	6074	22.6%
Chemical	1999Dec.-2000Nov.	Applicant	101	2221	4.5%
		Examiner	420	1799	23.3%
	2000Dec-2001Nov.	Applicant	115	1924	6.0%
		Examiner	431	1835	23.5%
Others	1999Dec.-2000Nov.	Applicant	249	5215	4.8%
		Examiner	1599	7571	21.1%
	2000Dec-2001Nov.	Applicant	327	4543	7.2%
		Examiner	1451	6362	22.8%

Appendix 2 Citation lag dependence of the number of total applicant citations by group of cited patent documents