RIETI Discussion Paper Series 15-E-096

Cognitive Distances in Prior Art Search by the Triadic Patent Offices: Empirical evidence from international search reports

WADA Tetsuo
Gakushuin University

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 15-E-096
August 2015

Cognitive Distances in Prior Art Search by the Triadic Patent Offices: Empirical evidence from international search reports[1]

WADA Tetsuo
Gakushuin University, Faculty of Economics

Abstract

Despite large numbers of empirical studies being conducted on examiner patent citations, few have scrutinized the cognitive limitations of officials at patent offices in searching for prior art to add patent citations during patent prosecution. This research takes advantage of the longitudinal gap between international search reports (ISRs) required by the Patent Cooperation Treaty (PCT) and subsequent examination procedures in national phases. It inspects whether several kinds of distances actually affect the probability that a piece of prior art is caught at the time of ISRs, which is much earlier than national phase examinations. Based on triadic PCT applications for all of the triadic patent offices (European Patent Office (EPO), United States Patent and Trademark Office (USPTO), and Japan Patent Office (JPO)) between 2002 and 2005 and their citations made by the triadic offices, evidence shows that geographical distances negatively affect the probability of prior patents being caught in ISRs, while a lag of prior art positively affects the probability. Also, the technological complexity of an application negatively affects the probability, whereas the size of forward citations of prior art affects it positively. These results show the existence of cognitive restrictions borne by officials at the patent offices, and suggest issues for designing work sharing by patent offices, in that the duplication of search costs exists only where search horizons of patent offices overlap each other.

Keywords: Examiner citations, Patent Cooperation Treaty (PCT), Triadic patents, International search reports (ISR), International patent families
JEL classification: K29, O33, O34

[1]This study is conducted as a part of the Project "Research on Innovation Process and its Institutional Infrastructure" undertaken at Research Institute of Economy, Trade and Industry (RIETI). The author is grateful for helpful comments and suggestions by the project leader Professor Sadao Nagaoka (Tokyo Keizai University), Professor Alfons Palangkaraya and Professor Beth Webster (Swinburne University), Profeesor Dietmar Harhoff (Max Planck Institute), Discussion Paper seminar participants at RIETI, and audience at the Asia-Pacific Innovation Conference 2014 and the 15th International Society of Scientometrics and Informetrics Conference. This research has also benefited from the collaborative project supported by RISTEX/JST with Professor Setsuko Asami (Tokyo University of Science) and Professor Yoshimi Okada (Hitotsubashi University), where a research focus is placed on the quality comparison between triadic patent offices. Professor Asami first thought of the comparison of search quality between PTOs at an aggregated level, based on ISRs. The author also acknowledges the support from the MEXT/JSPS (Grant #22330122) for the analyses of citations and firm boundaries.

1. Introduction

Patent citations have been widely utilized as empirical tools for studies on patent systems, particularly for such issues as economic value and knowledge flows (Trajtenberg, 1990, Hall et al.,2005; Jaffe et al., 1993). Although pioneering work did not distinguish examiner citations from applicant citations, subsequent studies have examined whether examiner citations are different from applicant citations. One of the studies on the subject was conducted by Alcacer and Gittleman (2006), who showed the similarity between examiner citations and inventor citations with respect to geographical distance in particular. While follow-up studies have further compared examiner citations and applicant citations with respect to other aspects such as the relationship with renewal rates (Hegde and Sampat, 2009) and the probability of use for rejections (Cotropia, Lemley, and Sampat, 2013), few have analysed how patent offices are influenced by several kinds of obstacles that can affect cognitive horizons during prior art search. Since examiners (and searchers working for a patent office) can never be perfect in conducting prior art search, the types and extent of the obstacles should be a part of the policy design parameters.

For example, recent development in the Patent Prosecution Highway (PPH) programs allows a patent office to utilize previous search and examination work conducted in an earlier prosecution process at a participating patent office of a PPH program, provided that patent applications are made in the two countries with the same priority date, and with corresponding (i.e., substantially the same set of) claims. The idea behind this program is to save the duplication of search and examination costs at the participating patent offices by sharing information gathered for the same invention. However, cost savings depend on whether or not the different patent offices actually make efforts in duplicate. If each of patent offices has very different set of advantages in technological knowledge that exists only locally, then examinations at the two offices are not duplicates. In order to design an international work-sharing plan between offices, the types and extent of obstacles of prior art search should be scrutinized to reduce duplicate search costs (or to obtain more complete search by overlapping search) by collaboration of their search processes. Little are known to date about such obstacles against searchers and examiners. It would be useful as a first step to define and test several possible “distances” for search when patent offices conduct search for prior arts.

One of the reasons why there have been no large-scale studies on the types and extent of obstacles against search by patent offices is the lack of measurement of the obstacles against search. This study utilizes International Search Reports (ISRs) as a basis for measuring such search obstacles of the triadic patent offices, and tests how officials are bounded by "distances," including geographical distances as well as similar kinds of cognitive obstacles against prior art search, without relying on comparison with applicant citations. In conducting the analyses, we consider applicants' self-selection, since applicants from the U.S. and Japan can choose the European Patent Office as their search agency, where the EPO has reputation for its quality in prior art search (therefore applicants who seek stringent search may choose the EPO ex ante).

2. Background and prior literature

Following the pioneering work of measuring the effect of knowledge spillover through patenting data (Jaffe 1986) and of measuring the value of patents through patent citations (Trajtenberg, 1990), the measurement of knowledge diffusion by patent citations was established by Jaffe and Trajtenberg (Jaffe et al. 1993, 1999). They have found that knowledge diffusion is geographically localized, assuming that patent citations show the traces of knowledge transmission. While survey results confirm that patent citations indicate knowledge flow with considerable noise (Jaffe et al. 2000, Duguet and MacGarvie 2005), there remains criticism against the method of measuring knowledge flow by patent citations, in that patent citations are often unrelated with knowledge transfer between inventors. This is partly because patent citations include examiner citations, and also because even applicant citations are sometimes added by attorneys on behalf of inventors in preparation for patent prosecution.

Since examiners are not inventing, and since inventors' perception of prior knowledge has been the central question for innovation research, one of the research agenda following the literature above is how "noisy" examiner citations are compared to applicant citations. Comparisons between examiner citations and applicant citations thus followed. Alcacer and Gittleman (2006) find that examiner citations and applicant citations have similar distributions in terms of the geographical distance between citing and cited patents. Their results based on the U.S. patent data partly contradicted with those from the EPO data, while examiner citations in the EU are found to be also bound by geographical distance (Criscuolo and Verspagen 2008). The advantage of examiner citations for economic research has been assessed in other meaning, such as in measuring the value of patent better than applicant/inventor citations (Hegde and Sampat, 2009). Further detailed comparison between applicant citations and examination citations has revealed that examiners do not rely on applicant-submitted information prior art (Cotropia, Lemley, and Sampat, 2013). Hoever, these studies did not analyse examiner citations independently.

Examiners are professionals to search for prior arts. Yet, their cognitive capabilities cannot be perfect. Recent micro-level studies on examiners' experience level and their granting behaviour (Lemley and Sampat, 2012; Frakes and Wasserman, 2014) as well as studies on examiner citations (Cotropia, Lemley, and Sampat, 2013) acknowledge the limitations of examiners. Except the small number of studies, economics literature has not analysed the extent that examiners are bounded in cognitive capabilities. A related line of research has been done at Melbourne (Jensen et al., 2005; Webster et al., 2007; Palangkaraya et al., 2011; Webster et al., 2014), which compared the results of patent grants from the triadic offices of the U.S., the EU, and Japan and have concluded that patent offices are biased toward local applicants (and against foreign applicants) in terms of patent grants. While it might be true that differential grant rates against foreign applicants can be caused by the "prejudiced" examinations in each office, examiners' cognitive bias (i.e, local advantage in technological knowledge) may contribute to the seemingly differential rates of patent grants as well. A remaining question is how we can measure examiners' cognitive limitations.

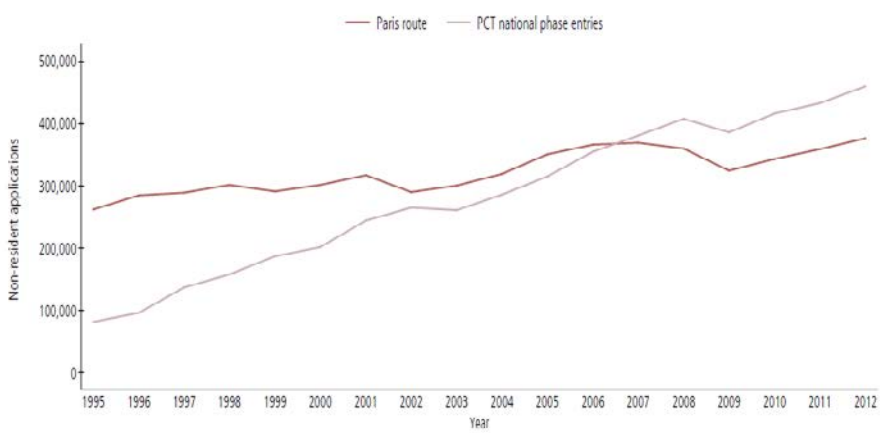
Most of the existing studies above utilize patent citation data from a single country or from two regions at most. Each data set of examiner citations in a country show only the results of a single patent office. If, however, we combine multi-regional citation data and consolidate the cited patents as well as citing patents through international patent families, in principle, we obtain a way to measure the difference between regions with respect to the same criterion: a particular cited patent family is indeed cited by the same patent family in multiple regions. Put differently, since examiner citations over different regions show the traces of examination outcomes in each region, how examiners behaved in terms of citing prior art can be tracked back. Depending on the assumptions of citation criteria in region, the difference of citing behaviour can provide the information on the cognitive limits of examiners. In the following section, this methodology is to be further explained.

3. The methodology: PCT and ISR as the basis of empirical measurement

The measurement of the search obstacle for examiners itself is the obstacle for research on examiners and searchers at the patent office. This paper proposes and implements a method of measuring the search obstacles, namely binding conditions on search capability, of the triadic patent offices by focusing on International Search Reports, or ISRs, issued by different International Search Authorities, or ISAs, specifically the patent offices in Europe, the U.S. and Japan, according to the PCT (Patent Cooperation Treaty).

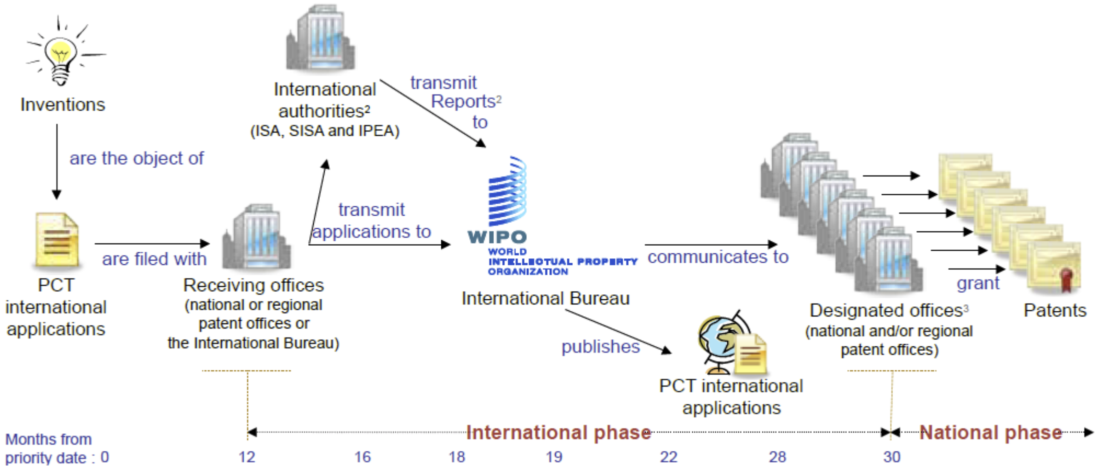
Before explaining the details of citation-level methodology, it should be noted that PCT applications are increasingly important for applicants who seek patent protections internationally, and that PCT applications should receive more attentions from economic research. The number of PCT national phase entries from abroad has already surpassed the number of non-resident applications via the Paris Convention route worldwide (World Intellectual Property Organization, 2014, p.48). In other words, the PCT is now the *main route* for international applications, whereas there have been few empirical studies on the PCT system. Given that the triadic offices (the European Patent Office, the United States Patent and Trademark Office, and the Japan Patent Office) received the vast majority of PCT applications before mid-2000s, it is reasonable to limit our samples to those PCT applications made to and examined at all of the three offices, at least until 2005.

Figure 1. Non-resident PCT and Paris Convention route entries into national phase worldwide (replicated from WIPO, 2014, p.48)



A PCT application received at an office will be given an ISR by an ISA at the time of international publication of the application. Under the PCT, "an applicant must file an application with a receiving office and choose an international searching authority to provide an international search report and a written opinion on the potential patentability of the invention (WIPO, 2014)." An ISR contains a list of prior arts, and the set of prior arts becomes a part of citations. ISRs are issued under the common search criterion established by the WIPO under the PCT system. "The applicant generally has at least 30 months from the filing (priority) date to decide whether to enter the national phase in the countries or regions in which protection is sought" (WIPO, 2014). The guideline at the WIPO applies to every ISA when issuing ISRs, whereas applicants in some countries are allowed to choose ISAs. The same criterion for prior art search is applied over different patent offices, while national phase examinations do not have such standardized rules. We can distinguish those cited patents, which are added in the national phase by the *designated offices* (or "*DO-cited*" patents), from those cited patents caught as citations at the earlier time of ISR ("*ISR-cited patents*").

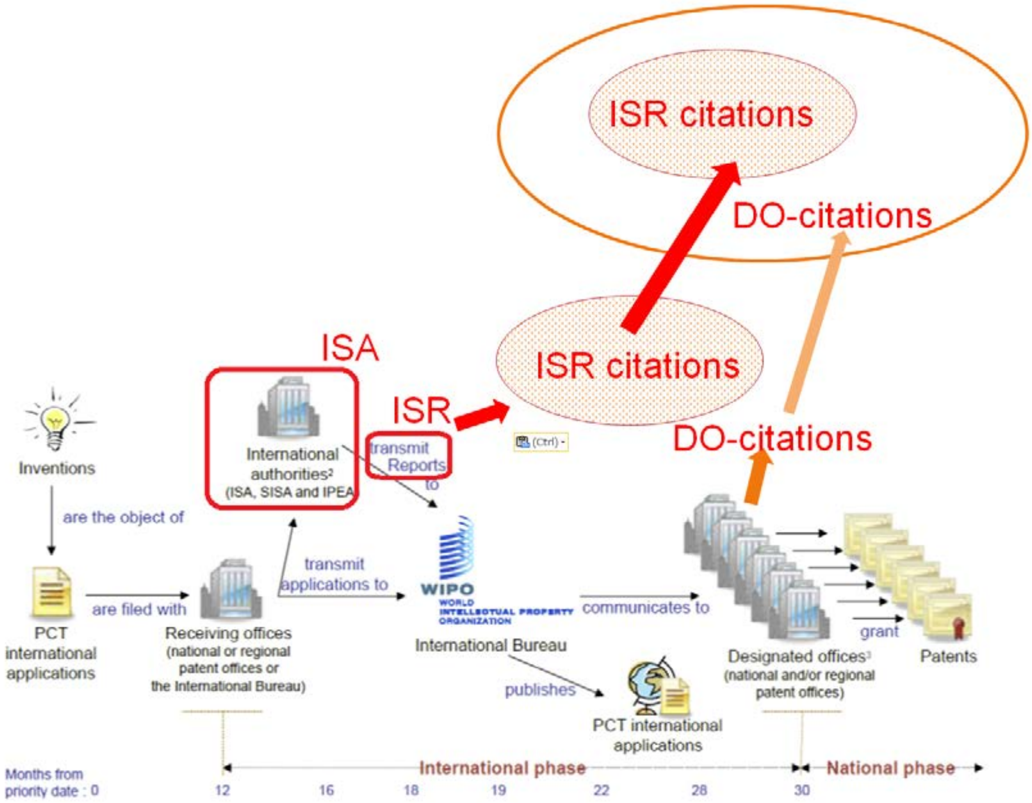
Figure 2. PCT procedure (replicated from WIPO, 2014, p.13)



As is seen in the Figure 2, there are time differences between ISRs and national phase examinations, implying the existence of lag between ISR-citations and DO-citations on the average in national phase. It allows a quasi-equal testing ground for search completeness of ISRs. While ISRs are issued at an early stage, more searches are conducted in national offices later. Since knowledge is geographically localized (Jaffe et al. 1993;1999), and knowledge diffusion takes time, additional time between

ISRs and national phase search facilitates more complete search in the later stage. We limit our samples to those PCT applications that are examined at all of the three triadic offices, meaning that localized knowledge in any of these areas at the time of ISRs is more likely to be caught by the offices at the national phase in a less localized way.

Figure 3. the dependent variable *found_in_ISR:* a binary dependent, representing the probability of a DO-citation or an ISR-citation to have been already included in the set of ISR-citations (modified on Figure 2. above)



Following the logic above, we retrospectively define the probability of every cited patent for a PCT application (union set of ISR-cited patents and DO-cited patents), consolidated and identified at the INPADOC family level, to have been already caught in the ISR of the originating PCT application (whether or not it was included in the ISR-cited patents). Taking this probability (a binary variable *found_in_ISR*, empirically) as the dependent variable, we implement PROBIT analyses at the INPADOC family level with explanatory variables representing the various "distances" between citing and cited patents, including technological complexity of originating applications, and other related indicators. Simply put, we assume that every DO-cited patent for a PCT application should have been cited in its ISR if every citer and cited pair is consolidated at the INPADOC family level, and if examiners (or searchers for an ISA) are unbounded in cognitive capability.

Several caveats should be mentioned on the methodology. First, applicants' (inventors') citations are excluded from the analysis, because the primary objective is to evaluate the determinant of search completeness by the ISAs. However, when an applicant is relatively capable in searching for prior art, its disclosure ex ante might affect the quality of search by a patent office. In this paper, additional analyses are conducted in order to consider the self-selection of the U.S. and Japanese applicants to choose the EPO as their ISA, since the EPO has high reputation of examination standard and therefore applicants with higher capability from the U.S. and Japan may choose the EPO as the ISA. This control methodology is not perfect, however.

Second, if a relevant prior art was missed at the time of ISR, it is assumed that one of the DOs can cite it. In reality, DO-citations are added according to different standards in different regions. Since the U.S. patent system do not provide citation category (such as "X" and "Y") information, we have not been able to apply the same standard of rejection for a cited piece of prior art. Also, DOs can never be perfect since they are also bounded by cognitive limits. Citations made by post-grant oppositions are included, but post-grant litigations are not. Thus, the union set of DO-citations is only an approximation of the quasi-complete search made possible ex post. On the other hand, DOs may cite a prior art in response to an applicant action such as amendment of claims, divisional applications, or continuations. Although ISAs are supposed to cite prior arts that are reasonably expected to be relevant in subsequent change of claims, not all prior arts that are triggered by ex post amendment of claims can be searched ex ante. Our basic assumption of "every DO-cited patent for a PCT application should have been cited in its ISR if ISAs are perfect in cognitive capabilities" may be violated if amendment of claims is too drastic.

Third, although actual ISR search is sometimes outsourced to non-PTO agencies, we consider ISRs as a basis of evaluating PTOs, since they are issued under the name of the patent offices, not private search agencies. Only citations made by the triadic offices are considered in the current analyses. That is, search completeness made possible by non-triadic offices is not considered in the analysis.

Fourth, since PATSAT, our primary data source, records non-patent literature in non-standardized formats, we could not consolidate the same non-patent literature across different records. For this reason, we employ patent citations only at this time. Also, US citations are not complete as well on PATSTAT. In particular, citations for rejected applications are not recorded on PATSTAT. Although it is usually possible to retrieve citations data from the Public PAIR database for rejected applications filed after 2001, we are still in the process of obtaining the data and have not been able to incorporate the data.

4. Hypotheses

Since ISR searchers (examiners/searchers for patent offices) are affected by cognitive obstacles from various "distances," we hypothesize that a prior patent (that was found in ISR or national phase) is more likely to be found in ISR when "distances" are less problematic, i.e., if:

H1) a relevant prior patent is closer in geography (less geographical distance),

H2) a prior patent is older (more knowledge diffusion time),

H3) a prior patent is from the same applicants (less organizational distance),

H4) a prior patent has more number of forward citations (more knowledge diffusion beforehand), and

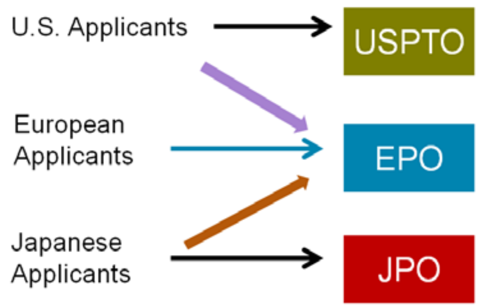
H5) an application for which an ISR is issued has less scope, less number of claims, less number of inventors, and less number of international family (less complexity against search).

In addition, we consider the possibilities that applicants' self-selection of ISAs affects the outcome variable. As is shown in the figure 4, PCT applicants from the U.S. and Japan are allowed to choose the EPO as their ISA[1]. On the contrary, applicants from the EU region are not allowed to choose the USPTO or the JPO as their ISA. Given the high reputation of the EPO for its quality in prior art search, applicants from the U.S. and Japan may self-select if they seek stringent search at the EPO. We therefore take the switching behaviour on PCT applications for ISRs as one of the factors for ISR

[1] In addition, recent agreement between the USPTO and the JPO allows applicants from the U.S. can choose the JPO as their ISA, starting 2015.

completeness, and use instrument variables for the ISA-switch (a binary "*ISA-changed*").

Figure 4. Selection of ISA from the U.S. and Japan as the Receiving Office (R.O.)



5. The Data source

The empirical domain of analysis is the triadic patent applications through PCT, with their earliest priority date within its international family between 2002 and 2005. Triadic PCT patent applications are defined here as INPADOC families that contain all of EPO, USPTO and JPO applications recorded on EPO's PATSTAT database, with only one "WO (PCT)" application in a family. It means that a single PCT application initiates international phase for all applications in a family. The number of international families for the analysis is 97,828. Although international applications to and from China and Korea has increased dramatically in the last ten years, the triadic patent offices of the EPO, the USPTO and the JPO represented the vast majority before 2005, which is our observation period.

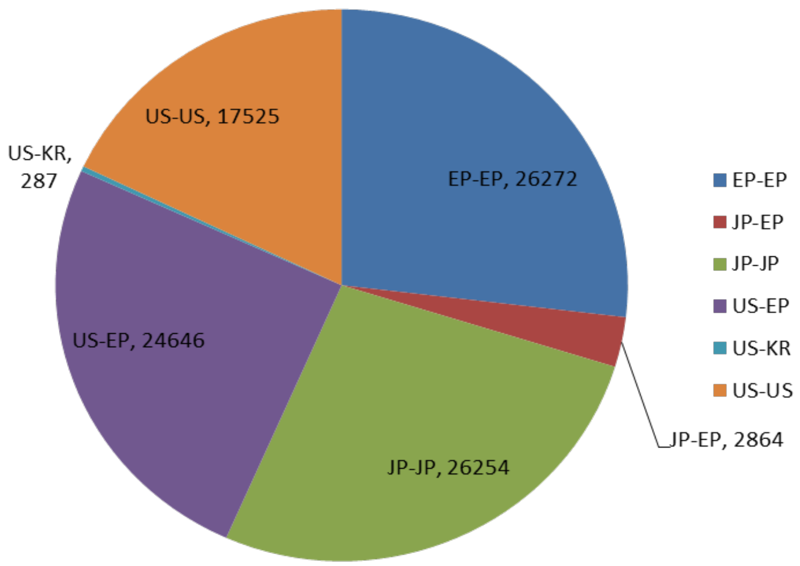
EPO PATSTAT (2013 OCT version) is used, and INPADOC family is the unit of analysis. Therefore, the accuracy of international families depends entirely on the INPADOC family table on PATSTAT. Citation data also comes from PATSTAT (2013 OCT), while JPO citation data is augmented by Seiri-Hyojunka data (JPO's standardized patent prosecution data). Applicant identifiers are consolidated by the EEE-PPAT database developed by ECOOM (Du Plessis et al., 2009; Magerman et al., 2009; and Peeter et al., 2009).

As was stated before, U.S. citations data are not complete as well on PATSTAT, since citations for rejected applications are not recorded on PATSTAT. Even after the publication rule change in the U.S. in 2001, published applications are not recorded on

PATSTAT if an application was abandoned (possibly due to rejection). The lack of the US citations for rejected applications may affect the result of the analysis, but this has not been verified yet.

Figure 5. Composition of triadic PCT applications, priority year 2002-2005

(JP-EP means the JPO being the receiving Office and the EPO being the ISA.)



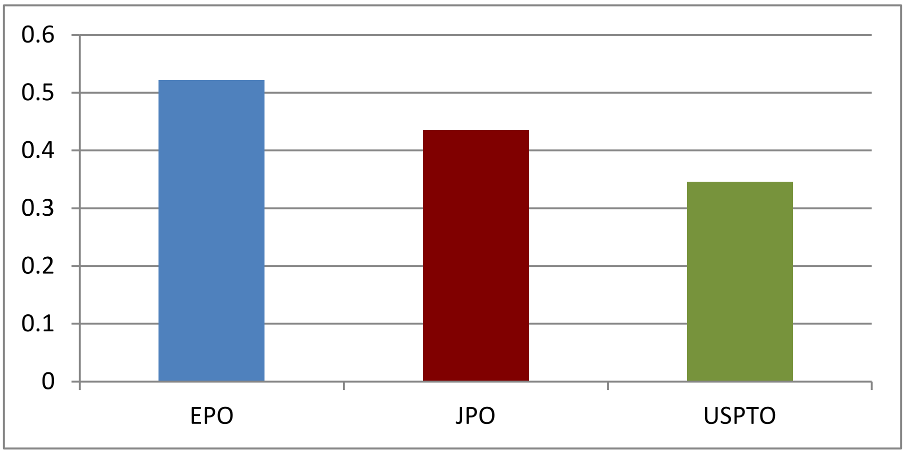
Based on the data set described, applications from the EPO area are shown to represent more than a quarter of the entire sample, as seen in Figure 5. In the figure, "JP-EP" stands for the JPO as Receiving Office and the EPO as ISA. "US-EP" stands for the USPTO as Receiving Office and the EPO as ISA. Applicants from the EPO area are not allowed to choose ISAs. On the other hand, applicants from the U.S. are allowed to choose ISAs from the EPO, the IP Australia, the Korean Intellectual Property Office (KIPO), the Rospatent (Russian Patent Office), etc. In fact, more than half of PCT applications from the U.S. choose the EPO as ISA, while 0.7% choose KIPO[2] as ISA. Applicants from Japan are allowed to choose from the JPO and the EPO as their

[2] This small share of KIPO is partly due to the fact that our samples are limited to EPO-USPTO-JPO triadic applications between 2002-2005. Recently, KIPO has the third largest share in ISR publications. In 2013, the EPO remained the most selected ISA, with 37.7% of all ISRs issued, followed by the JPO at 20.7%, KIPO at 14.8%, and USPTO at 8.1%. (WIPO, 2014, pp.68-69).

ISA, but approximately only one tenths of Japanese PCT applications have chosen the EPO as ISA.

Behind the selection of the ISAs, there are differences in the reputation of completeness of search reports between the triadic offices, i.e., the EPO has the highest reputation. A simple comparison between the average of *found_in_ISR*, according to the three ISAs, shown in Figure 6, is consistent with the reputation. Given that the EPO has good reputation, and given further that applicants from the U.S. or Japan can choose the ISA, we expect that self-selection by applicants influences the outcome variable, *found_in_ISR*. This is partly because applicants with inventions of higher economic value or with higher capability would spend more cost for prior art search by themselves, so that they would find more prior art before submitting applications. Furthermore, highly capable applicants may seek more stringent search at its early stage, so that rejection in later stage, i.e., national phase, will not occur. Indeed, there is evidence[3] revealing that applicants know the EPO produces higher quality ISRs in general, although its cost is higher than the cost charged by other offices. In order to account for the self-selection, we hypothesize that the more experienced and capable an applicant in the U.S. or in Japan is in technological innovation, it is more likely for the applicant to choose the EPO as its ISA.

Figure 6. Simple average of the dependent variable *found_in_ISR*, according to ISA



[3] "Which PCT International Search Authority (ISA) should I use?" Web page on "inovia.com," browsed on July 12th, 2014. http://info.inovia.com/2013/09/which-pct-international-search-authority-isa-should-i-use/#sthash.91J9FjDM.dpuf

6. Variables

We employ several categories of explanatory variables, representing each of hypotheses above, in PROBIT analyses taking the probability of a cited patent being caught in the previous ISR as the binary dependent variable (“*found_in_ISR”)*. The unit of analysis is a pair of citing and cited international families, both consolidated at INPADOC family level.

For H1, three variables of *euro_cited* (cited family has its 1st priority, i.e., the earliest date, in EPC countries within a family, derived from tls201 and tls219 tables of PATSTAT), *us_cited* (cited family has its 1st priority in the U.S.), and *jp_cited* (cited family has its 1st priority in Japan) are defined. When a cited family has its origin in the same region where ISR is issued, the ISA of the region is expected to have geographical advantage over the relevant technology. Expected sign is positive for each region, e.g., positive *jp_cited* coefficients for applications originating from Japan.

For H2, citation lag between the 1st priority of a citing family and that of a cited family is defined as *fam_cite_lag* (derived from tls201 and tls219 tables of PATSTAT). The longer the lag is, the easier the prior art is to be found at the time of ISR. Therefore, its expected sign is positive.

For H3, *self* is defined as a binary variable, taking the value of one if one of patents in a cited family and one of patents in a citing family belongs to the same applicant, based on PATSTAT (tls207) combined with EEE-PPAT, using “L2” id. Patent office will find it easier to locate prior relevant art within the same applicant. Therefore, its expected sign is positive.

For H4, *fwd_cite_of_the_cited* is defined and obtained from PATSTAT (tls217) as the number of forward examiner citations, counted at publication level (but consolidated at family level), and made out to the cited patent family. When a prior art has been already cited by many patents, patent offices will find it easier. Therefore, its expected sign is positive.

For H5, we first use scope indicators. *IPC4_count* is the total net count of IPC subclasses (4-digit IPC, derived from tls209) assigned in a citing INPADOC family. Since patent classification of an application may change during prosecution process both in international phase and in national phase, we include all IPC subclasses to capture the breadth of a family. The number of claims of a patent is correlated with the complexity of the technological content. As an indicator of the number of claims, we obtain *publn_claims_max_tls211*, which is the maximum number of claims registered on PATSTAT (tls211 table) in a citing INPADOC family. We do not simply rely on claims data from a single office such as from the EPO, since an application can be modified during its prosecution internationally. We also employ *invt_nr*, the maximum number of inventors in an application included in a citing INPADOC family, from PATSTAT (tls207). The size of international family, *family_size*, is a count variable of applications in different countries in a citing INPADOC family (tls211/219). Because all of the complexity measurement works negatively for prior art search by patent offices, those expected signs are all negative.

In addition to the variables above, which are used to test hypotheses directly, we define three variables to address self-selection of ISAs by applicants. The first two represent the potential of the applicant. The first of the two is *total_count*, which is the number of total applications that an applicant has made, taken from EEE-PPAT. The second one is *applicant_avg_cited*, which is the number of average forward citations that an applicant has received for an application, calculated by PATSTAT (tls212) and EEE-PPAT. Both are supposed to represent the experience level of the applicant, and are used as instrument variables for instrumented PROBIT on the variable *ISA_CHANGED.* This binary variable *ISA_CHANGED* indicates that the U.S. and Japanese applicants choose the EPO as their ISA (the EPO can be chosen from the U.S. and Japanese applicants, but not vice versa). This information can be obtained for PCT applications on PATSTAT, since the citation table tls212 has a field on "citation origin" where "ISR" is shown for PCT applications. Since first application country (RO) in a family is available from tls201, switching from RO to a different ISA can be coded. The correlation coefficient between *ISA_CHANGED* and the dependent variable *found_in_ISR* is low at around 0.03.

Control variables for originating areas, which are *JP_app* and *US_app* (applications from Japan and the U.S., respectively), are used. Technology class is controlled by thirty-five WIPO technology classification dummies, taking the last class as the reference class.

7. Estimation results

The result shown in the Model 1 of Table 1 employs samples only from the EPO regions. H1 is supported from the positive sign of *euro_cited* and negative signs of *us_cited* and *jp_cited*. Likewise, H2, H3, H4, and H5 are all supported in this model, except that the number of inventors has a positive sign for the coefficient, contrary to the expectation from H5. Model 1-2 further limits the samples to those from the EPO region and non-self citations only for a robustness check. The results are unchanged from the model 1. Model 1-3 employs all triadic samples from the EPO, USPTO, and JPO regions, with *JP_app* and *US_app* as applicants' region controls. It implies that the EPO samples are set as the reference. The coefficients of the two region controls have negative and significant signs, meaning that ISRs prepared by the USPTO and the JPO or applications are disadvantaged than those by the EPO on the average. The binary variable *ISA_CHANGED* is added to indicate if the U.S. and Japanese applicants choose the EPO as their ISA. The coefficient of the *ISA_CHANGED* is positive and significant, meaning that switched ISA from the USPTO or from the JPO to the EPO has made an ISR more complete. Results on other variables are mostly unchanged from the Model 1 and model 1-2, except that the coefficient for the number of inventors has lost significance. The coefficient for *jp_cited* has shifted from a negative to a positive sign, but this is due to the pooled samples. It means that prior arts from the JPO area are easier to be found by the EPO on the average, if compared to all samples. The results from the Model 1 and 1-2 clearly shows that prior arts from the JPO area are more difficult to be found by the EPO on the average, if compared to the EPO samples only.

Model 2-1 uses applications from the U.S. only, and all of the results are consistent with the hypotheses. Model 2-2 also focuses on the U.S., and limits the citation data to non-self citations only for robustness checks, while employing two instrument variables on the variable *ISA_CHANGED* through instrumented Probit. The results is almost unchanged. The only exception is that the coefficient of *IPC4_count* lost significance.

Model 3-1 uses samples from Japan only in order to examine the locality of knowledge in Japan. As is expected in H1, *jp_cited* has a positive and significant sign, whereas *us_cited* has negative and significant sign. Other variables show similar results with the Model 1 and 2 and are consistent with hypotheses, except that *self* indicates a negative sign and shows an insignificant coefficient for the number of inventors. For Japanese applications, the coefficient for *ISA_changed* lost the significance in the Model 3-2, suggesting that the advantage provided by the ISA change from JPO to EPO is due to the applicants' self-selection. However, this effect is not observed for the U.S. applications in the Model 2-2.

Some of the results related with the WIPO thirty-five technology classes are noteworthy. The coefficients for the classes 14, 15 and 16 have positive signs consistently (15 and 16 in particular). The WIPO field classifications number 14 is for "Organic fine chemistry, 15 for "Biotechnology," and 16 for "Pharmaceuticals" (See Appendix). Those technological classes are known for discrete technologies, and a patent has high economic value on the average, compared to complex technologies. Since applicants conduct relatively complete search before filing applications in those classes, prior arts on the ISRs are thought to be relatively complete.

Table 1. PROBIT analyses on the probability of ISR coverage; dep. var.=*found_in_ISR*

****<0.001 ***<0.005 **<0.01 *<0.05 Robust standard errors are in the parentheses (clustering on citing family).

Model & sample	Model 1 (EP_app only)	Model 1-2 (EP app & non-self only)	Model 1-3 (EP, US and JP apps all pooled)
method	Probit	Probit	Probit
ISA_CHANGED			0.3096426**** (0.0066579)
euro_cited	0.207075**** (0.016451)	0.2196004**** (0.017633)	0.1419984**** (0.0080393)
us_cited	-0.0456823** (0.016363)	-0.0523266*** (0.017525)	-0.0620007**** (0.0078305)
jp_cited	-0.457428**** (0.0169483)	-0.4633691**** (0.0181224)	0.0393056**** (0.0082601)
fam_cite_lag	0.003277**** (0.0003825)	0.0025981**** (0.0003906)	0.0030127**** (0.000212)
self	0.0635864**** (0.0103293)		0.2091817**** (0.0047187)
fwd_cite_of_the_cited	0.0000356**** (0.00000785)	0.0000321**** (0.00000783)	0.0000359**** (0.00000321)
IPC4_count	-0.0116202**** (0.0026942)	-0.0106372**** (0.0028234)	-0.0165033**** (0.0013614)
publn_claims_max_tls211	-0.0142597**** (0.0004492)	-0.014192**** (0.0004704)	-0.0080901**** (0.0001942)
invt_nr	0.0071474*** (0.0023555)	0.009076**** (0.0025359)	0.0000932 (0.0011831)
family_size	-0.0064288**** (0.0012222)	-0.0074571**** (0.0013424)	-0.006626**** (0.0007439)
JP_app			-0.0667862**** (0.0069462)
US_app			-0.2808785**** (0.0072769)
tech_field1	0.1446844*** (0.0447274)	0.1499931*** (0.0462607)	0.1444009**** (0.0271212)
tech_field2	0.1826226**** (0.0500017)	0.1890902**** (0.0511893)	0.0698306* (0.0274406)
tech_field3	0.1663785*** (0.0529874)	0.1589762*** (0.054606)	0.0329559 (0.0286252)
tech_field4	-0.0212283 (0.0485813)	-0.0195436 (0.0498217)	-0.0591332* (0.0277592)
tech_field5	0.0396309 (0.0627064)	0.0327979 (0.0648507)	0.0154652 (0.0335855)
tech_field6	0.0762131 (0.0476598)	0.0771954 (0.048991)	-0.0270184 (0.0274349)
tech_field7	0.1203062 (0.1757166)	0.1112099 (0.1794752)	0.202257**** (0.0420252)
tech_field8	0.2287141**** (0.0542056)	0.2308582**** (0.0562138)	0.0823158*** (0.0278554)

tech_field9	0.3632714**** (0.0504529)	0.3818553**** (0.0526016)	0.2023969**** (0.0283424)
tech_field10	0.0869068 (0.0446431)	0.0865379 (0.0461657)	0.073116** (0.0275033)
tech_field11	0.573764**** (0.0684509)	0.5695354**** (0.0717487)	0.452349**** (0.0364449)
tech_field12	0.0436221 (0.0597322)	0.0402991 (0.0614565)	0.0208029 (0.0350831)
tech_field13	0.2541932**** (0.0457902)	0.2437701**** (0.0473539)	0.1369031**** (0.0270368)
tech_field14	0.6062752**** (0.0436612)	0.6254198**** (0.0456541)	0.5058929**** (0.0275027)
tech_field15	0.7705994**** (0.0507952)	0.7553919**** (0.0539247)	0.5908729**** (0.0289889)
tech_field16	0.7391506**** (0.0444679)	0.7145057**** (0.0466877)	0.5942508**** (0.0274072)
tech_field17	0.4043448**** (0.0447423)	0.4241324**** (0.0466949)	0.2743913**** (0.0278738)
tech_field18	0.7592531**** (0.0791369)	0.7672647**** (0.0839064)	0.4946623**** (0.0391235)
tech_field19	0.5697184**** (0.045845)	0.569902**** (0.0480957)	0.339838**** (0.0286062)
tech_field20	0.3680992**** (0.0490161)	0.3547626**** (0.0506313)	0.1884243**** (0.0290885)
tech_field21	0.3057705**** (0.0547929)	0.2878575**** (0.0573025)	0.2135224**** (0.0302265)
tech_field22	0.0804922 (0.1395774)	0.1227397 (0.1436416)	-0.0830575 (0.0682282)
tech_field23	0.1434451*** (0.048968)	0.1493542*** (0.0507068)	0.1351532**** (0.0297479)
tech_field24	0.1409634* (0.0554432)	0.152254** (0.0573317)	0.1049673*** (0.0324254)
tech_field25	0.0837243 (0.0491715)	0.0882502 (0.0509684)	0.0083187 (0.029498)
tech_field26	0.0327776 (0.0480379)	0.0359581 (0.0496099)	-0.0087836 (0.0295359)
tech_field27	0.1589274**** (0.0450578)	0.1717163**** (0.0466351)	0.0817932*** (0.0287871)
tech_field28	0.3042161**** (0.0528322)	0.3184216**** (0.0556424)	0.2153852**** (0.0298522)
tech_field29	0.3330521**** (0.0527045)	0.3400129**** (0.054884)	0.1929519**** (0.0304772)
tech_field30	0.2264897**** (0.0598906)	0.241079**** (0.0615502)	0.1368849**** (0.0342744)
tech_field31	0.0094726 (0.0447409)	0.0004919 (0.0465451)	0.0446885 (0.0287091)
tech_field32	0.0457021 (0.0436223)	0.0444188 (0.0451479)	0.0033423 (0.0282497)
tech_field33	0.0576639 (0.0678101)	0.0663571 (0.0692349)	0.0616716 (0.0388799)
tech_field34	0.1700641** (0.0608972)	0.1883208*** (0.063318)	0.1239305**** (0.034291)

tech_field35	(reference)	(reference)	(reference)
_cons	-0.1400582*** (0.0460341)	-0.1352025** (0.0479984)	-0.2124146**** (0.028279)
Log pseudo-likelihood	-158417	-135935	-661846
N	249307	214766	1,031,127
# of clustered citing families	26078	25318	97125

Table 1 (continued). PROBIT analyses on the probability of ISR coverage; dep. var.=*found_in_ISR*

Model 2-2 and 3-2 use "*total_count*" and "*applicant_avg_cited*" as instruments for "ISA_CHANGED."

****<0.001 ***<0.005 **<0.01 *<0.05 Robust standard errors are in the parentheses (clustering on citing family).

Model & sample	Model 2-1 (US app only)	Model 2-2 (US app & non-self only)	Model 3-1 (JP app only)	Model 3-2 (JP app & non-self only)
method	Probit	IV Probit	Probit	IV Probit
ISA_CHANGED	0.380766**** (0.0074961)	1.35421**** (0.3121828)	0.2758815**** (0.0169662)	0.010949 (0.1314657)
euro_cited	0.1776262**** (0.0120059)	0.148418**** (0.025389)	-0.031025 (0.0160179)	0.0203394 (0.0174625)
us_cited	0.050351**** (0.0114757)	0.0777813**** (0.015989)	-0.3377195**** (0.0155267)	-0.2974986**** (0.0169034)
jp_cited	-0.4295359**** (0.0121628)	-0.3751167**** (0.0427645)	0.8054234**** (0.0151802)	0.8367819**** (0.0175193)
fam_cite_lag	0.0046464**** (0.000329)	0.0026492**** (0.0005495)	0.0023379**** (0.0004175)	0.0005303 (0.0004425)
self	0.1123806**** (0.0076398)		-0.1759722**** (0.0082345)	
fwd_cite_of_the_cited	0.0000573**** (0.00000437)	0.0000551**** (0.00000526)	-0.00000566 (0.00000781)	-0.00000566 (0.00000799)
IPC4_count	-0.0215867**** (0.0022476)	0.0099131 (0.0110926)	-0.0176023**** (0.002381)	-0.0170435**** (0.0026306)
publn_claims_max_tls211	-0.0094453**** (0.0002733)	-0.0081833**** (0.0010323)	-0.0029271**** (0.0003468)	-0.0033284**** (0.0004149)
invt_nr	-0.0058672*** (0.0018111)	-0.0089979*** (0.0026536)	-0.0007108 (0.002112)	0.0008906 (0.0023144)
family_size	-0.0053694**** (0.0011327)	-0.0138593**** (0.0024961)	-0.0142835**** (0.0021553)	-0.0091501*** (0.0032126)
tech_field1	0.2428436**** (0.0486241)	-0.0146199 (0.104442)	0.0558508 (0.052299)	0.0819208 (0.0592823)
tech_field2	0.1705577*** (0.0493154)	-0.0698391 (0.1006609)	-0.0889339 (0.052323)	-0.0670638 (0.0592019)
tech_field3	0.0491601 (0.0491757)	-0.1330145 (0.0858773)	-0.0105879 (0.0550479)	0.0107328 (0.0615837)
tech_field4	-0.051275 (0.047647)	-0.2180253** (0.0796503)	0.0289418 (0.0556962)	0.0609505 (0.0621146)
tech_field5	0.0793476 (0.0562173)	-0.1366787 (0.1099556)	-0.016325 (0.0639327)	0.0388485 (0.071582)
tech_field6	-0.0329754 (0.0475002)	-0.1670512* (0.0761786)	0.0185922 (0.0542866)	0.0406843 (0.0609839)
tech_field7	0.1822923*** (0.0611697)	0.3286845** (0.1177346)	0.301815**** (0.0778456)	0.2877073*** (0.0844477)
tech_field8	0.1597099*** (0.0485588)	-0.1225467 (0.1126863)	-0.015146 (0.0533185)	0.0062975 (0.0599288)
tech_field9	0.2777179**** (0.0500796)	-0.0917514 (0.1392015)	0.0778071 (0.053753)	0.1100494 (0.0601185)

tech_field10	0.1435841*** (0.0483644)	-0.1696981 (0.1150714)	0.0614314 (0.0539854)	0.0919096 (0.0601537)
tech_field11	0.4360864**** (0.0566237)	0.2004486 (0.1047643)	0.4875324**** (0.0778804)	0.4982929**** (0.0862923)
tech_field12	0.092297 (0.0577402)	-0.244581* (0.1239686)	-0.0532708 (0.0671496)	-0.0637358 (0.0721385)
tech_field13	0.1019339* (0.046934)	-0.1875399 (0.1069902)	0.1091174* (0.0545724)	0.1280565* (0.0607982)
tech_field14	0.4952246**** (0.0483835)	0.1383903 (0.1460763)	0.5494072**** (0.0566347)	0.6247042**** (0.0645541)
tech_field15	0.5244111**** (0.0489719)	0.3890508**** (0.0903527)	0.6803185**** (0.0597871)	0.714162**** (0.0675115)
tech_field16	0.4855671**** (0.0475861)	0.2599136* (0.1045442)	0.7830664**** (0.0563774)	0.8116951**** (0.0635088)
tech_field17	0.3765414**** (0.0499197)	-0.1800641 (0.2036643)	0.124376* (0.0545345)	0.1169751 (0.0610249)
tech_field18	0.4386131**** (0.0637165)	0.1800015 (0.1398228)	0.3238745**** (0.0699398)	0.3073437**** (0.0791576)
tech_field19	0.3128908**** (0.049961)	-0.1536425 (0.1698294)	0.1735532*** (0.057277)	0.199093*** (0.0632013)
tech_field20	0.2963152**** (0.053857)	0.0091053 (0.1195746)	0.0366946 (0.0545363)	0.0480973 (0.0611276)
tech_field21	0.3192602**** (0.0516942)	-0.0179226 (0.1381289)	0.0727353 (0.057422)	0.0690492 (0.0641752)
tech_field22	-0.0679129 (0.0983615)	-0.5369288** (0.1984481)	-0.1661115 (0.1859442)	-0.0750412 (0.2190826)
tech_field23	0.2447045**** (0.0510565)	-0.045901 (0.1199372)	0.0652941 (0.0579753)	0.0724272 (0.0653014)
tech_field24	0.1936245*** (0.0572679)	0.0493307 (0.0954071)	0.073069 (0.0609101)	0.1169597 (0.0672215)
tech_field25	0.0765811 (0.0517362)	-0.1360526 (0.0942277)	-0.1090352 (0.0570776)	-0.0940606 (0.0639862)
tech_field26	0.1624198*** (0.0531509)	-0.0905513 (0.1068294)	-0.127283* (0.0556927)	-0.103975 (0.0627597)
tech_field27	0.1813037*** (0.055923)	0.0724606 (0.0936175)	0.0177349 (0.0555072)	0.0953296 (0.0656141)
tech_field28	0.2468054**** (0.0519663)	-0.1588447 (0.1502847)	0.114324* (0.0566256)	0.1558597* (0.0629745)
tech_field29	0.2274386**** (0.0550812)	-0.1564595 (0.1523844)	0.0931103 (0.056731)	0.0832719 (0.0631814)
tech_field30	0.2590443**** (0.0636966)	0.0802455 (0.1164088)	-0.0191918 (0.0608549)	-0.0317799 (0.0671536)
tech_field31	0.1797807*** (0.0540232)	-0.0533882 (0.1014106)	0.0473454 (0.0552925)	0.0693065 (0.061321)
tech_field32	0.057358 (0.0546367)	-0.0956397 (0.0949013)	0.0048358 (0.0555261)	0.0423921 (0.062192)
tech_field33	0.1509299* (0.0664652)	-0.0736027 (0.1271391)	-0.0436442 (0.0680619)	-0.045587 (0.0741097)
tech_field34	0.1494443** (0.0570063)	-0.1447077 (0.117844)	0.1162017 (0.0657658)	0.1356205 (0.0727767)
tech_field35	(reference)	(reference)	(reference)	(reference)

_cons	-0.4960139**** (0.0489832)	-0.8014725**** (0.1324103)	-0.6243859**** (0.0553713)	-0.6797762**** (0.0615108)
Log pseudo-likelihood	-276135.	-467661	-197428	-186380
N	455830	363328	325990	264805
# of clustered citing families	41074	38066	28973	28099

8. Discussion and further development

Overall results are consistent with the hypotheses, suggesting that examiners (and searchers working for the PTOs) are bound by various kinds of "distances," including technological complexity of applications. These are not very surprising, but are supported by the novel methodology for the first time. Examiners (unlike inventors) are required to find prior art by law, but they are naturally bound by informational horizons they have. Most prior studies using examiner citations do not incorporate these informational obstacles born by examiners, and this present study has simply established a methodology to determine the existence of barriers. As was stated in the review of prior literature, prior studies on the difference of examination outcomes between patent offices (Jensen et al. 2005; Webster et al., 2007, 2014) have not explicitly considered them. Taking the cost of prior art search into the grant rate comparison would be a fruitful way of extending the research envelope. However, as is explained in the methodology and data sections, there are several limitations to be solved. In particular, the data on the U.S. needs filtering on citation categories, and augmentation on rejected (abandoned) applications. Also, the results with instrument variables suggest the self-selection is working, but is evident for the Japanese samples only. Further scrutiny is needed with updated data on more attributes of applicants and applications.

These results have policy implications, since Patent Prosecution Highways (PPH) rely on outcomes from previous patent offices. Given that knowledge is locally concentrated due to agglomeration economies, a local patent office may have an advantage over other distant patent offices to find relevant prior knowledge existent locally. This is also likely since local examiners are educated and employed locally and have access to up-to-date information in local language. In other words, the physical distance between the location of an invention and the location of its relevant prior art is not independent from the probability of the prior art to be found by examiners (and searchers employed or contracted for patent offices). If we try to evaluate the merit by combining the work done by more than one patent office, an efficiency question depends on how distant patent offices make duplicate effort with each other. Put differently, in order to justify a system of physically dispersed patent offices on the planet, rather than a unitary single patent office that searches and examines patent applications worldwide, we need to know how complementary the offices are in terms of searching capabilities. This paper shows the initial step toward the policy question.

Table 2. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
found_in_isr	1057671	0.387615	0.487206	0	1
ISA_CHANGED	1057671	0.276951	0.447492	0	1
euro_cited	1057671	0.192777	0.39448	0	1
us_cited	1057671	0.434949	0.495751	0	1
jp_cited	1057671	0.357832	0.479362	0	1
fam_cite_lag	1042360	9.420731	8.568765	-5	50
self	1057671	0.140923	0.347942	0	1
fwd_cite_of_the_cited	1057671	76.10043	419.7362	1	21950
IPC4_count	1057671	3.35843	1.919123	1	25
publn_claims_max_tls211	1057671	19.1512	16.08566	0	296
invt_nr	1057671	3.099957	2.137287	1	39
family_size	1057671	7.833797	3.451029	4	41
total_count	1057671	9898.959	24339.66	0	115208
applicant_avg_cited	1001720	0.832080	1.836592	0	84.75
JP_app	1057671	0.313767	0.464023	0	1
US_app	1057671	0.44353	0.496801	0	1

Table 3. Variables

found_in_isr	A binary variable, indicating a cited patent being caught in the previous ISR
ISA_CHANGED	ISA changed to EPO (PATSTAT)
euro_cited	cited patent has its 1st priority in EPC countries (PATSTAT)
us_cited	cited patent has its 1st priority in the US (PATSTAT)
jp_cited	cited patent has its 1st priority in Japan (PATSTAT)
fam_cite_lag	citation lag between the 1st priority of a citing family and that of a cited family (PATSTAT)
Self	examiner citation within the same applicant (PATSTAT&EEE-PPAT)
fwd_cite_of_the_cited	# of forward examiner citations (sum in a family) in PATSTAT
IPC4_count	the total net count of IPC subclasses (4-digit IPC) assigned in an INPADOC family (PATSTAT)
publn_claims_max_tls211	# of claims (maximum in an INPADOC family on PATSTAT tls 211 table)
invt_nr	# of inventors (PATSTAT)
family_size	# of applications in the same international family (PATSTAT)
total_count	# of total application that an applicant has made (EEE-PPAT)
applicant_avg_cited	# of average forward citations that an applicant has received per its patent (PATSTAT&EEE-PPAT)
JP_app	JPO as RO (PATSTAT)
US_app	USPTO as RO (PATSTAT)

Table 4. Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	found_in_isr	1														
2	ISA_CHANGED	0.031	1													
3	euro_cited	0.0763	−0.0631	1												
4	us_cited	−0.0721	0.1867	−0.4016	1											
5	jp_cited	0.016	−0.1431	−0.3544	−0.6462	1										
6	fam_cite_lag	0.0187	−0.0021	0.07	0.0259	−0.0746	1									
7	self	0.0778	0.006	0.0354	−0.0214	0.0139	−0.1865	1								
8	fwd_cite_of the cited	−0.0008	0.0246	−0.0414	0.1222	−0.0903	0.0216	−0.0019	1							
9	IPC4_count	−0.0131	−0.0296	0.0021	0.002	−0.0024	−0.0244	0.0106	0.0148	1						
10	publn_claims_ max_tls211	−0.123	0.1023	−0.0611	0.0972	−0.0527	−0.0455	−0.0179	0.0324	0.1018	1					
11	invt_nr	0.0201	0.0202	0.015	0.0086	−0.0186	−0.0344	0.0559	0.0106	0.0981	0.0744	1				
12	family_size	0.022	0.0558	0.114	0.0557	−0.1495	0.0415	0.0471	−0.0015	0.0916	0.0832	0.1262	1			
13	total_count	−0.0147	−0.106	−0.0907	−0.1094	0.1781	−0.0983	0.0221	−0.0028	−0.0571	−0.0135	0.0077	−0.1658	1		
14	applicant_avg _cited	0.0268	0.1515	0.0227	0.1262	−0.1475	−0.0269	0.0203	0.0372	0.0651	0.059	0.0569	0.1793	−0.159	1	
15	JP_app	0.0141	−0.3031	−0.1598	−0.2679	0.4046	−0.0593	0.05	−0.0364	−0.0115	−0.0896	−0.0145	−0.2677	0.4326	−0.2631	1
16	US_app	−0.0644	0.5958	−0.1046	0.343	−0.2697	−0.0092	−0.0398	0.0575	0.0082	0.197	0.0099	0.0653	−0.2782	0.2602	−0.6049

References

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.

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

Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. Research Policy, 37(10), 1892-1908.

Du Plessis, M., Van Looy, B., Song, X & Magerman, T. (2009). Data Production Methods for Harmonized Patent Indicators: Assignee sector allocation. EUROSTAT Working Paper and Studies, Luxembourg.

Duguet, E. & M. MacGarvie. (2005) How well do patent citations measure flows of technology? Evidence from French innovation surveys. Economics of Innovation and New Technology 14.5: 375-393.

Frakes, M. & M. Wasserman (2014) Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents?: Evidence from Micro-Level Application Data, NBER Working Paper No. 20337.

Jaffe, A. (1986). Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. The American Economic Review, 76(5), 984-1001.

Jaffe, A., & Trajtenberg, M. (1999). International knowledge flows: evidence from patent citations. *Economics of Innovation and New Technology* 8, 105-136.

Jaffe, A., M. Trajtenberg, & M. Fogarty. (2000) Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *American Economic Review* 90.2: 215-218.

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

Jensen, P.H., Palangkaraya, A. & Webster, E. (2005) Disharmony in international patent office decisions. *Federal Circuit Bar Journal*, 15, 679.

Lemley, M. A., & Sampat, B. (2012). Examiner characteristics and patent office outcomes. Review of Economics and Statistics, 94(3), 817-827.

Magerman T, Grouwels J., Song X. & Van Looy B. (2009). Data Production Methods for Harmonized Patent Indicators: Patentee Name Harmonization." EUROSTAT Working Paper and Studies, Luxembourg.

Palangkaraya, Alfons, Elizabeth Webster, and Paul H. Jensen. "Misclassification between patent offices: evidence from a matched sample of patent applications." Review of Economics and Statistics 93.3 (2011): 1063-1075.

Peeters B., Song X., Callaert J., Grouwels J., Van Looy B. (2009). Harmonizing harmonized patentee names: an exploratory assessment of top patentees. EUROSTAT working paper and Studies, Luxembourg.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. The Rand Journal of Economics, 172-187.

Wada, T., & Asami, S. (2014). Quality comparison of International Search Reports (ISRs) by selectable International Search Authorities (ISAs) under the Patent Cooperation Treaty (PCT) system, *a paper presentation at the 2014 International Workshop on Patent System Design for Innovation*, Hitotsubashi University, Tokyo, Japan.

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

Webster, E, Palangkaraya, A & Jensen, P.H. (2007) Characteristics of international patent application outcomes. *Economics Letters* 95, 362-368.

World Intellectual Property Organization. (2014). Patent Cooperation Treaty Yearly Review: The International Patent System. *WIPO Economics and Statistics Series*.

Appendix. WIPO technology fields

Field_number	Field_name
1	Electrical machinery, apparatus, energy
2	Audio-visual technology
3	Telecommunications
4	Digital communication
5	Basic communication processes
6	Computer technology
7	IT methods for management
8	Semiconductors
9	Optics
10	Measurement
11	Analysis of biological materials
12	Control
13	Medical technology
14	Organic fine chemistry
15	Biotechnology
16	Pharmaceuticals
17	Macromolecular chemistry, polymers
18	Food chemistry
19	Basic materials chemistry
20	Materials, metallurgy
21	Surface technology, coating
22	Micro-structural and nano-technology
23	Chemical engineering
24	Environmental technology
25	Handling
26	Machine tools
27	Engines, pumps, turbines
28	Textile and paper machines
29	Other special machines
30	Thermal processes and apparatus
31	Mechanical elements
32	Transport
33	Furniture, games
34	Other consumer goods
35	Civil engineering

WIPO, “World Intellectual Property Indicators 2011,” p.181