INGENIO WORKING PAPER SERIES

Examiner amendments to applications to the european patent office: Procedures, knowledge bases and country specificities

Joaquín M. Azagra-Caro and Elena M. Tur

Working Paper Nº 2014-06

INGENIO [CSIC-UPV] Ciudad Politécnica de la Innovación | Edif 8E 4º | Camino de Vera s/n | 46022 Valencia
tel +34 963 877 048 | fax +34 963 877 991 | innco@ingenio.upv.es

Examiner amendments to applications to the european patent office: Procedures, knowledge bases and country specificities

Joaquín M. Azagra-Caro, Elena M. Tur[1]

INGENIO (CSIC-UPV), Valencia (Spain)

Abstract: The geography of knowledge flows has shown that the probability of a patent applicant rather than the examiner originating a citation depends on differences between citing and cited countries. How the characteristics of the citing country affect that probability has received less attention. Using European Patent Office (EPO) data of over 3,500,000 citations (1997-2007), we find that the probability of applicant citation is higher as national economic and scientific strengths increase, if applicants and examiners come from the same country and if the country belongs to EPO. This 'country club' effect is comparable to that found for US Patent and Trademark Office.

Keywords: Citations, knowledge flows, knowledge spillovers, national biases
JEL Codes: O30, O33, O34

1 Introduction

The geography of innovation makes extensive use of backward citations in patents to measure knowledge flows (Jaffe et al., 1993). Several works emphasize the importance of distinguishing the origins of citations because, in theory, citations inserted by patent

[1] The research was carried out with funding from project GV/2012/018 of the Valencian Regional Government. A previous version of the paper was presented at the T2S 2013 Conference. Carlos Benito and África Villanueva helped with brainstorming. Davide Consoli participated in a related paper that inspired this one. Jordi Molas-Gallart revised a previous version. Ester Planells offered technical assistance. Davide Lingua clarified issues related to the information in Patstat.

examiners are likely to be less localized than applicant citations. United States Patent and Trademark Office (USPTO) data mostly tend to confirm this for the US case (Thompson, 2006) although there are some differences for some specific measures of distance (Alcácer & Gittelman, 2006). European data confirm it for some European countries (Criscuolo & Verspagen, 2008), but not some regions with low absorptive capacity (Azagra-Caro et al., 2009). These studies focus on the match or distance between citing and cited country. However there is another geographic concern that has been largely unexplored, i.e. what are the characteristics of the citing country? Do patent examiners add more citations to patent applications from specific countries?

This is a relevant question because the answer might reveal underlying economic forces that are subject to policy influence, or uncover individual questionable examiner practices. There is some evidence suggesting that, for the USPTO, geographic origin of the applicant matters, e.g. US examiners add more citations to foreign applications (Alcácer et al., 2009).[2] However, there is a lack of research on a similar 'club effect' in the case of the European Patent Office (EPO). This is unfortunate because the EPO is frequently used as a benchmark against the USPTO, and is considered one of the highest quality patent systems due to its rigorous granting process and flexibility applied to later stages in a patent's life (Saint-George & van Pottelsberghe, 2013). This paper focuses on the EPO. By comparing with the USPTO, we should be able to identify whether there is a symmetrical geographical effect, namely whether EPO patent examiners are more likely to add citations to foreign applications.

This implies a need first to define 'foreign' in the context of the EPO. The question of national differences in patent examiner versus patent applicant inserted citations is

[2] Even within a European NUTS2 region, NUTS3 regional per capita income is negatively associated with (Spanish) national examiner citations and positively associated with inventor citations (Azagra-Caro et al., 2011).

especially interesting in the case of the EPO because it is an international patent office. In national patent offices, applicants seek protection in one country only, which may or may not be their own, so examiners can clearly differentiate between domestic and foreign applicants. In the EPO, the distinction between domestic and foreign applicants is blurred, since grants provide protection for inventions in many countries, and examiners can include people of different nationalities. National patent offices may judge foreign applicants according to criteria that are not the same as those of the applicant's country's patent office whereas the EPO applies international criteria, agreed upon by the signatory countries to the European Patent Convention, i.e. members of the European Patent Organization (EPOrg), to judge applications from different countries. Therefore, our research questions are: Do EPO examiners add extra citations to applications from countries outside the EPOrg? And do EPO examiners add extra citations to applications from countries other than their own?

In an analysis of over 3,500,000 backward citations in years 1990-2007 drawn from EPO data, we found substantial differences in the share of examiner citations across applicant countries (Source: Patstat, edition October 2012). In our view, these variations deserve explanation.

To do this, we isolate the influence of many sub-national variables, i.e. characteristics of citations, patents and applicants. We already know that differences in the citation shares of applicants and examiners vary according to type of patent and type of applicant (Alcácer et al., 2009; Azagra-Caro et al., 2011). This paper contributes to fine-grained analysis of the phases of the patenting process and types of citations. This should increase our understanding of the effect of origin of citations in the EPO. Are they associated with granted patents (which seems not to be the case in the USPTO)? Are they associated with the scientific knowledge base of the patent?

To put this research in context, we discuss some of these sub-national aspects in detail to understand citation processes at the EPO (section 2). We present the data (section 3) and the results (section 4), and finally we conclude (section 5).

2 Citation processes in the EPO

2.1 Identifying the distinction between applicant and inventor citations at the EPO

Information on differences between applicant and examiner citations in EPO applications is provided in the search report (Criscuolo & Verspagen, 2008). Examiners produce search reports and include citations to previous art. They divide citations into 'categories'. The categorization is specific to the search report formats of the EPO, PCT, and many other authorities.[3] The patent examiner decides which citations in the application text are relevant to the examination and therefore should be included in the search report (category D). The examiner also adds (usually a majority of) citations, in several categories (X, Y, A, etc.). Table 1 provides figures for the distribution in the sample studied here.

{Table 1 around here}

The biggest block of citations refers to the state-of-the-art without challenging the patentability of the invention (category A). The second biggest block (based on number of citations) indicates possible challenges to patentability, and so is qualitatively more important because it reduces the probability of a patent being granted (categories X and Y). The third biggest block is citations included by the applicant (category D).

[3] However, not all authorities use this format (the US does not unless it is a PCT search report, then the USPTO has to provide the data required by the PCT), so they do not provide the categories. Moreover, if citations are added by examiners (or in the opposition procedure) following the search report, these are not categorized.

Some studies explain the meaning of these categories X, Y and A in more detail (e.g. Meyer, 2000). For the purposes of the present study, what matters is the share of category D citations over total citations which is used as a proxy for the share of applicant citations (Criscuolo & Verspagen, 2008).

2.2 EPO procedures and search reports

Criscuolo & Verspagen (2008) analyze citations from the search report. They do not provide details on which search report and there are three types of documents that are referred to as search reports: EPO search reports, examination reports, and international search reports. A precise understanding of citation phases could help clarify this issue. Figure 1 represents the process.

{Figure 1 around here}

Citations originate in several phases of the patent's life. Not all patent offices require completion of all phases.[4] Not every phase generates applicant and examiner citations simultaneously. For direct EPO applications, there are two relevant phases. First, the EPO search procedure which ends with publication of an EPO search report. The main purpose of the search report is to document whether the application fulfills the requirements of novelty and non-obviousness. It is not binding but it signals to the applicant the likelihood of the patent being granted, which informs the decision about whether to pay the extra fees required to continue the process.

The second phase consists of the substantive examination. Examiners assess industry applications in addition to novelty and non-obviousness. This phase ends with the publication of an examination report and grant or denial of the patent.

[4] An overview can be found in file REFI Statistics 2014.xlsx, open to the public in: http://documents.epo.org/projects/babylon/rawdata.nsf/0/a655ef0be2534bdbc1257991002894cf?OpenDocument [2/4/2014].

The EPO allows indirect applications through the PCT procedure. The PCT authorizes some patent offices to be International Search Authorities and run a unified application protocol, valid in all signatory countries. In this case, there is a phase that produces applicant and examiner citations which precedes the first two phases described above, and results in an international search report published by the PCT. After publication of the international search report, if the applicant designates the EPO as 'region' of protection, the application enters the 'European regional phase' and becomes a Euro-PCT application, which then enters the two phases described above.

3 Model, data and variables

We estimate the following model:

$$\Pr(appcit_{ijklt}) = f(\alpha X_{it}, \beta X_{jt}, \gamma X_{kt}, \delta X_{lt}) + \varepsilon_{ijklt} \qquad (1)$$

where *appcit* is equal to 1 if the citation is inserted by the applicant and 0 if inserted by the examiner. The probability varies according to the characteristics of the citation *i*, the patent *j*, the applicant *k* and the applicant country *l*. The year of the patent application *t*, is lagged two periods for national economic and research and development (R&D) characteristics to prevent endogeneity.

Data on patents and citations come from Patstat (October 2012 edition). Figure 2 provides a visualization of the process. We selected patents where the publication authority was the EPO –almost 2.5 million. After removing those with missing or unreliable information for application year[5] and technology class (represented by the

[5] We excluded the period after 2008 because Patstat does not include citation categories after that year. We also excluded 2008 given the decrease in the number of applications which is probably due to statistical not real reasons (delay induced by the grant process).

International Patent Classification IPC), and those without citations,[6] we were left with 2 million patents.

{Figure 2 around here}

Those patents contained over 12 million citations (Figure 3). Patstat classifies them into origin types, i.e. the moment in the examination process when the citation was inserted. There are ten types of origins (coded 0-9), but only some are relevant for this study, i.e. those indicating that either patent applicant or examiner could have inserted the citation (see section 2.2 for further details): origins coded 0 (citations introduced during search), 2 (citations introduced during examination) and 5 (citations from the International Search Report). They represent most (82%) of the citations.

Patstat differentiates who inserted the citation by classifying citations with origins 0, 2 and 5 into several categories. Categories (coded with single letters, A, X, Y, etc.), refer to the relevance of prior art to invalidate claims of novelty. Criscuolo & Verspagen (2008) call category D 'applicant citations' and sum the other categories as 'examiner citations'. We follow this method (see section 2.1 for further details).

{Figure 3 around here}

In the estimations, the number of observations is not the number of citations for two reasons (Figure 4). First, duplicates are created if the patent has more than one applicant. We deal with this econometrically by weighting the observations by the inverse number of applicants. Second, we match Patstat to other databases on national characteristics that do not have full information for all countries and years. The sample includes over 3.6

[6] Some PCT patents have comments in Patstat such as "See references of WO 0046271A1". The data for those references are accessible through the international application identifier field, and citations in that format were included in the study through the use of an algorithm of our own creation. However, due to an inconsistency in the Patstat database, some PCT patents do not have international identification numbers which meant we were unable to cross reference their data with their cited documents.

million observations. The proportion of D-citations in the total is our dependent variable, computable for over 7 million citations.

{Figure 4 around here}

The evolution of applicant citation shares has been declining in the period of observation, 1997-2007 (Figure 5), following the pattern detected by Criscuolo & Verspagen (2008: Fig. 1) for the period 1985-2000.

{Figure 5 around here}

National variation in the sample is clear (Figure 6), with many core European Union states showing the highest applicant citation shares. A club effect is already apparent since the highest examiner citation shares correspond to countries that do not belong to the EPOrg: US, Japan, Korea, China and Australia.[7]

{Figure 6 around here}

In Europe, a pattern favoring the main Western countries emerges (Figure 7): applicant citation shares are larger for Belgium, Netherlands, Switzerland, Austria, Germany and France while examiner citation shares are larger for peripheral countries such as Spain, Poland, Italy and Greece. This suggests large applicant citation shares are related to economic wealth. There are exceptions for both sides: East European, lower income countries such as Slovenia and Hungary stand out for applicant citation shares and higher income countries such as Finland stand out for examiner citation shares. Hence, other factors matter, and are accounted for in the econometric estimations.

{Figure 7 around here}

Table 2 provides information on the econometric model variables.

---

[7] See full list of member states: http://www.epo.org/about-us/organisation/member-states.html [2/5/2014].

{Table 2 around here}

The dependent variable is a dummy that takes the value 1 if the citation comes from the examiner. A logit model is appropriate for this kind of data.

4 Results

Table 3 presents the estimations. Column 1 includes the specification of Equation (1) with citation and patent characteristics only; the remaining columns include the variables progressively.

{Table 3 around here}

4.1 Citation and patent characteristics

The results for the sub-national variables are consistent across estimations. We begin with EPO procedural aspects, where we expect a strong 'early intervention' effect: in shorter application procedures and in earlier phases applicants should include more citations than examiners. Citations are coded to indicate whether the origin is a Euro-PCT (not a direct EPO) application, and whether it is the European search report or the examiner report (rather than the international search report). The coefficient of "Euro-PCT" is negative and significant, indicating that this longer procedure leads to higher numbers of examiner citations. The coefficient of "European search report" is negative and significant, implying that citations in this second phase are more likely to be associated with examiners than if there was an international search report in the first phase. The coefficient of "Examiner report" is also negative and significant and higher than the coefficient of "European search report", meaning that citations in this third phase are most likely to come from examiners. These results are all consistent with the explanations in section 2.2, and consequently EPO measures using pooled data will be biased towards finding more citations related to Euro-PCT and international search reports than if only applicant citations are used.

The sample includes applications and grants. If examiner citations imply that the application is not well documented and/or not very novel, we would expect lower applicant citation shares on applications unlikely to be granted. Patstat includes codes that allow identification of whether an application was granted or not. This is controlled for in the models by the dummy variable "Grant". The estimated coefficient is positive and significant. Hence, we can confirm a link between receiving relatively fewer examiner citations and having the patent granted. In part, this is intuitive. It becomes more interesting if we consider that, in the USPTO, this does not necessarily apply. In the USPTO, more experienced examiners, and examiners that systematically cite less prior art, are more likely to award patent grants (Lemley & Sampat, 2012). Moreover, USPTO examiners rarely use applicant citations to reject a grant (Cotropia et al., 2013). Hence, examiner citation shares are not associated with denial of a grant in the USPTO but they are in the EPO. This and other signs may indicate the superiority of the EPO patent system (Saint-George & van Pottelsberghe, 2013).

Citations can be to patent literature or non-patent literature. The study of knowledge flows through patent citations focuses on patent literature (Jaffe et al., 1993) and refers to difference between examiners and applicants (Thompson, 2006; Alcácer & Gittelman, 2006; Criscuolo & Verspagen, 2008). However, citations to non-patent literature are interesting because they signal science relatedness (Callaert et al., 2013) and their study is mandatory in the context of the EPO because they are more frequent in the EPO than in other patent offices (Michel & Bettels, 2001; Callaert et al., 2006). We test whether applicants are more likely than examiners to cite non-patent literature, extrapolating from US evidence that examiners rarely cite non-patent literature (Sampat, 2004). The positive and significant sign of "Non-patent literature" shows that this is the case. Applicants are probably more familiar with the fundamental knowledge base underpinning their inventions, while examiners are often engineers whose expertise is related more to parcels of applied knowledge. This large-scale sample provides evidence that use of pooled citation data most likely underestimates the importance of science relatedness for patenting compared to use of only applicant citations.

4.2 Applicant characteristics

The expected effect of type of organization on examiner citation shares is tricky. Firms are the main patentees and may be more familiar with the rules of the game, which in turn may make their applications less subject to insertion of examiner citations. However, universities and government labs are closer to the science base and also are often familiar with the relevant literature, which will also reduce examiner added citations. Serial academic invention is based on the quality of earlier patents (Lawson and Sterzi, 2013), which is another reason to expect lower examiner citation shares. Dummies for organizational type of the applicant (models 2-3) can be used to validate empirically what matters more. "Company only" is the benchmark. The positive, significant coefficients of "Government only" and "University only" indicate that these institutions generate more reliability than corporate patents.[8] The coefficients of "Individuals only" and "Hospital only" are negative and significant, which means that citations are less likely to originate in applicants than in the case of firms. Individuals may show lower citation shares because institutions facilitate settings where citing is more common practice, i.e. through sharing of references and codified knowledge. Examiner citation shares may be larger for

---

[8] The coefficient of "University only" is positive and significant in model 3 but not model 2, suggesting that the inclusion of national characteristics is important to estimate precisely the effect of the applicant's organizational type.

private hospitals because they do not have a tradition of patenting, and on patents related to clinical practice which are less related to science.

Patents co-applied for by more than one organizational type are evidence of technological cooperation between the partners. Among the different possible relationships, university and/or government-industry interaction has been shown to have potential benefits. Industry historically has been interested in ensuring that higher education institutions provide practical solutions to technical needs, even in low-tech regions (Ortega-Colomer, 2013). Firms get access to doctoral graduates through cooperation with universities (Garcia-Quevedo et al., 2012). For academics, establishing appropriate combinations of strong and weak ties with external actors, mainly firms, increases research output (Villanueva-Felez et al., 2013). If university and/or government-industry interaction are so useful, one would expect to find them linked to larger applicant citation shares, because they would originate better documented and more novel patent applications. Models 2-3 include dummies for types of organizational interactions (taking "Company only" as benchmark). University-company co-applications for patents are strongly associated with a higher probability of an applicant rather than the examiner including a citation. Somewhat surprisingly, government-company co-application for patents is negatively related to that probability. A possible reason might be that organizations in the category government have heterogeneous missions. Government labs with an industry orientation are more likely to engage in partnerships with firms that lead to patents, than labs with an academic orientation, and the government-company dummy captures this type of partnership. This double industry orientation receives a higher share of examiner citations. For other interactions

("Company-hospital", "University-government") the dependent variable does not change significantly.[9]

More experienced applicants may have learnt the rules of the game and know better how to lower the cost of reference search efforts. Therefore, one would expect a higher probability of applicant citations in patents from repeat patentees. The database allows for the inclusion of a variable based on total number of applications filed by the same applicant. In contrast to our expectations, the number of applicant citations decreases with the increase in the number of applications. Alcácer et al. (2009) found the same in the USPTO case. Their explanation is that large applicants prefer "broad patent portfolios, with relatively low value placed on any single invention" (p. 426). Alternatively, it might be that applicants include unrelated cites after the invention or omit relevant cites for strategic reasons (Breschi & Lissoni, 2005). Perhaps experienced applicants learn how to "cheat", and hide a higher number of relevant references. Both views need further investigation, and could be the subject of future research. However, the present evidence shows that the EPO and the USPTO are similar in this regard.

4.3 National characteristics

National characteristics have a variety of effects (model 3). Ideally, applicants use citations to claim the novelty of their invention, arguing improvements in the state of the art. Examiners use citations for the opposite reason: to claim that novelty is not so high given the existence of similar inventions or bodies of knowledge. This implies that patent applications where the proportion of applicant citations is high are probably patents that the examiner considers quite novel, because he/she does not need to refute many claims. Thus, larger, wealthier and scientifically stronger countries are more likely to create

[9] We dropped from the sample a few observations with triple institutional interactions because their scarce numbers provoked multicollinearity.

conditions favorable to the appearance of novelty. The variables GDP, per capita GDP and GERD intensity test this assumption. Their positive, significant coefficients provide evidence to support it. Hence, we observe that countries with these favorable endowments benefit from lower examiner citation shares. This corresponds with other evidence showing that these countries are more innovative (Furman et al., 2002).

The composition of R&D by source of funding has concerned some scholars. High shares of business funding of R&D are associated with higher innovative capacity measured by patents at the national level (Furman et al., 2002). This is associated with the applied orientation of national research and lower knowledge flows (Azagra-Caro & Consoli, 2014). Applied orientation and lower knowledge flows may be at odds with novelty because they imply narrower scope than basic research. Hence, we may expect a lower probability of an applicant originating a citation than the examiner, in countries with higher shares of business funding of R&D. The coefficient of the share of business funding variable in model 3 is negative and significant, supporting this expectation. Examiner citation shares are higher in patents from national contexts where the research orientation is towards more applied research.

Country block effects may also play a role in the model. Specifically, we are interested in whether there is a club effect similar to the one shown by Alcácer et al. (2009) in the USPTO case: US applicants receive fewer examiner citation shares than non-US ones. In our EPO sample, this club effect would not be strictly national since the EPO is international. Instead, we propose that such an effect might be visible for countries belonging to the EPOrg. This may be due to better knowledge of the ‘rules of the game’ among signatories to the EPOrg and may translate into larger examiner citation shares for non-EPOrg countries. In the model, the dummy is equal to 1 if the applicant country belongs to EPOrg, to capture this phenomenon. The estimation (positive and significant) verifies that there is a lower propensity for EPOrg member states to receive cites from the examiner. Hence, the EPO is similar to the USPTO: outsiders are less warmly received.

Having isolated a club effect, it is possible that the nationality of examiners might be influential. Collins & Wyatt (1988) detected national chauvinism in citations to non-patent literature in US genetics patents: "it appears that every country is its own best citer" (p.73). However, Meyer (2000) finds no signs of national chauvinism in nanotechnology patent applications to the USPTO from Swedish applicants, but finds evidence of cosmopolitanism – perhaps due to small country size. If we assume the presence of national chauvinism in examiner citation shares, it may show up in lower citations to the home countries of examiners. This may be due to a common cultural background that includes knowledge of similar references (including in the native language), or to deliberate misconduct. The positive, significant coefficient of the probability of an application being examined by an examiner from the same country as the patent applicant provides support for the national chauvinism assumption. Examiners are more likely to favor (in terms of adding fewer citations) patents from applicants of the same nationality as themselves.

5 Towards an interpretative framework

Previous research has not tackled in depth what the differences between applicant or examiner citations might mean. Early works were aimed at disentangling whether related indicators changed results about geographical spillovers (Thompson, 2006; Alcácer & Gittelman, 2006; Criscuolo & Verspagen, 2008). The studies by Alcácer et al. (2009) and Azagra-Caro et al. (2011) made advances by qualifying applicant citation shares as indicators of 'applicant search effort' and examiner citation shares as indicators of 'examiner trust' on the probability that a patent application is granted. Both works provide intuitive but speculative and arguable reasons for using those labels. The present paper does not aim to offer a definitive answer and interviews with key actors would result in more accurate information. However, the significant statistical associations found may help to support use of the labels 'applicant search effort' and 'examiner trust'.

It seems that high applicant citation shares are related to the possibility of getting the patent being granted which is what one would expect if applicant search efforts were rewarded. Being awarded a patent could be considered synonymous with examiner trust. Examiner citation shares are lower in the presence of a scientific knowledge base, which could be expected if examiners trust that basicness fulfils the legal requirement for novelty.[10] Examiner citation shares also decrease in applications from richer and more R&D intensive countries, which implies that these countries are responsible for more valid filings (indicating more applicant search effort) or can rely on positive reputation (indicating examiner trust).

In this framework, we would interpret the results as follows. Examiner trust depends positively on the science-relatedness of the knowledge base, and on national economic and research strengths. This seems compatible with the theory of well functioning patent systems. The desirability of other significant influences is more ambiguous. Examiner trust could also be positively related to country-block club effects and national chauvinism. These may be understood in terms of common rules of the game and common cultural background. However, examiner trust might also indicate that trust is a socially mediated construct, and that awarding a patent may be subject to certain inefficiencies. To summarize, these interpretations, or analogous ones if we use the label 'applicant search effort' instead of 'examiner trust', depend on how much we believe that

---

[10] This holds at different aggregation levels: the citation (non-patent literature), the institution (university, government, company-university) and the country (% of non-business funding of R&D).

the labels are adequate. This study provides quantitative support to previous conjectures in order to stimulate further discussion.

6 Conclusions

The literature on the geography of knowledge flows has shown that the probability of an applicant rather than the examiner originating a citation depends on differences between citing and cited countries (Thompson, 2006; Alcácer & Gittelman, 2006; Criscuolo & Verspagen, 2008).[11] Our contribution to this stream of literature is that the conditions of the citing country also matter to predict that probability. Our findings show that better national economic and scientific endowments increase applicant citation shares, whereas higher proportions of business funding of R&D foster examiner citation shares. Future research could test which group of determinants (citing country characteristics or citing-cited country differences) matter more.

Previous analyses of the characteristics of applicant versus examiner citation shares found differences across patent and applicant characteristics (Alcácer et al., 2009; Azagra-Caro et al., 2011). We show the presence of additional disparities across citation characteristics, namely procedural aspects of the patenting process and knowledge base of the patent. Our results for procedural aspects increase our understanding of the generation of citations in the various phases of the life of an EPO application. Our results for knowledge base suggest the importance of science to provide credibility to applications.

---

[11] This has led to critiques of the use of pooled backward citations to the extent of disregarding non-patent literature to measure university-industry collaboration (Chen et al., 2012), or justifying the use of applicant citations only (Acosta et al., 2013). Others have suggested that backward citations, whether added by examiners or applicants, are valid as indicators of technological relatedness (Barirani et al., 2013). Survey results indicate that reliance on non-patent literature as indicators of scientific flows depends on the type of citation, with applicant self-citations being the most reliable (Li et al., 2014).

The use of a sample based on EPO applications allowed comparison with earlier works exploiting USPTO evidence. It suggests that large applicant citation shares are more clearly associated with being awarded a patent by the EPO than the USPTO (Lemley & Sampat, 2009; Cotropia et al., 2011). It also signals that there are similar club effects, which favor EPOrg members at the EPO and US residents at the USPTO (Alcácer el al., 2009). Since the methods used by Alcácer et al. (2009) and those applied in this study differ, interpretation of this comparison should be cautious. A possible avenue of further inquiry could be designing an experiment to enable direct comparison between both data sources.

Another line of investigation could exploit the distinction between applicant and examinant citations to search for potential weaknesses in examiner practices (Collins & Wyatt, 1998; Meyer, 2000; Lemley & Sampat, 2009; Cotropia et al., 2011). This would mostly concern the individual level and our paper does not allow straightforward comparison. However, at the aggregate level, the present study offers large-scale quantitative evidence of national bias at the EPO. After controlling for differences in technology fields, national economic and R&D strength, EPO club effect, etc., we find that the nationality of examiners still conditions the proportion of applicant citations. This suggests the presence of a systematic bias that cannot be attributed to the other factors we control for. This bias may be due to limitations in the cultural knowledge bases of EPO examiners or to national chauvinism. Resolution of either would require improvements to the selection and/or training of EPO examiners. However, our analysis is limited by lack of data on the examiners of individual patent applications; we rely on national aggregates and indicators of the probabilities of finding an examiner from the same country. Qualitative research could provide additional support for this finding.

Overall, as previous research shows, we lack a definitive interpretation of the meaning of applicant/examiner citation shares. In the previous section, we supported the idea of higher values for applicant citation shares as an indicator of higher examiner trust. If this is the case, our results highlight the benefits of fostering the scientific knowledge base, public research, university-industry interaction and national economic and scientific strengths for promoting trust in technological knowledge production. This conclusion is tentative and is suggested to encourage more efforts dedicated to resolving this issue.

References

Acosta, M., D. Coronado, R. Marín, and P. Prats, "Factors affecting the diffusion of patented military technology in the field of weapons and ammunition," Scientometrics 94 (2013), 1–22.

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

Alcácer, J., M. Gittelman, and B.N. Sampat, "Applicant and examiner citations in U.S. patents: An overview and analysis," Research Policy 38 (2009), 415-427.

Azagra-Caro, J.M., and D. Consoli, "National determinants of knowledge flows: the moderating role of public-private cooperation," Journal of Technology Transfer (2014), forthcoming.

Azagra-Caro, J.M., I. Fernández de Lucio, F. Perruchas, and P. Mattsson, "What do patent examiner inserted citations indicate for a region with low absorptive capacity?" Scientometrics 80:2 (2009), 443–457.

Azagra-Caro, J.M., P. Mattsson, and F. Perruchas, "Smoothing the lies: The distinctive effects of patent characteristics on examiner and applicant citations," Journal of the American Society for Information Science and Technology 62:9 (2011), 1727-1740.

Barirani, A., B. Agard, and C. Beaudry, "Discovering and assessing fields of expertise in nanomedicine: a patent co-citation network perspective," Scientometrics 94 (2013), 1111–1136.

Breschi, S., and F. Lissoni, "Knowledge networks from patent data," In H. F. Moed, W. Glänzel and U. Schmoch (Eds.), Handbook of quantitative science and technology research (Netherlands: Kluwer Acad. Publ.. 2005).

Callaert, J., M. Pellens, and B. van Looy, "Sources of inspiration? Making sense of scientific references in patents," Scientometrics (2013), doi 10.1007/s11192-013-1073-x.

Callaert, J., B. van Looy, A. Verbeek, K. Debackere, and B. Thus, "Traces of Prior Art: An analysis of non-patent references found in patent documents," Scientometrics 69(1) (2006), 3-20.

Chen, C.M., H. Su, and P. Lee, "Does Non-Patent Reference Measure University-Industry Collaboration?" Proceedings of PICMET '12: Technology Management for Emerging Technologies (2012), 1049-1053.

Collins, P., and S. Wyatt, "Citations in patents to the basic research literature," Research Policy 17 (1988), 65-74.

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

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

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

Furman, J.L., M.E. Porter, and S. Stern, “The determinants of national innovative capacity,” Research Policy 31 (2002), 899–933.

Garcia-Quevedo, J., F. Mas-Verdú, and J. Polo-Otero, “Which firms want PhDs? An analysis of the determinants of the demand,” Higher Education 63:5 (2012), 607-620.

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

Lawson, C., and V. Sterzi, “The role of early-career factors in the formation of serial academic inventors,” Science and Public Policy (2013), doi: 10.1093/scipol/sct076.

Lemley, M., and B.N. Sampat, “Examiner Characteristics and Patent Office Outcomes,” Review of Economics and Statistics 94:3 (2012), 817-827.

Li, R., T. Chambers, Y. Ding, G. Zhang, and L. Meng, “Patent citation analysis: Calculating science linkage based on citing motivation,” Journal of the Association for Information Science and Technology (2014), doi: 10.1002/asi.23054.

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

Meyer, M., “Does science push technology? Patents citing scientific literature,” Research Policy 29 (2000), 409-434.

Michel, J., and B. Bettels, “Patent citation analysis. A closer look at the basic input data from patent search reports,” Scientometrics 51:1 (2001), 185-201.

Ortega-Colomer, F.J., “The evolution of the local role(s) of the university in a low-tech region,” International Journal of Technology Management & Sustainable Development 12:1 (2013), 71–87.

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

Saint-George, M., and B. van Pottelsberghe de la Potterie, “A quality index for patent systems,” Research Policy 42 (2013), 704– 719.

Sampat, B.. “Examining Patent Examination: An Analysis of Examiner and Applicant Generated Prior Art, Working Paper (2004), School of Public Policy, Georgia Institute of Technology.

Thompson, P. “Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations,” Review of Economics and Statistics 88:2 (2006), 383-388.

Villanueva-Felez, A., J. Molas-Gallart, and A. Escribá-Esteve, “Measuring Personal Networks and Their Relationship with Scientific Production,” Minerva 51 (2013), 465–483.

Figures

Figure 1 - Phases with either applicant or inventor citations in EPO patents

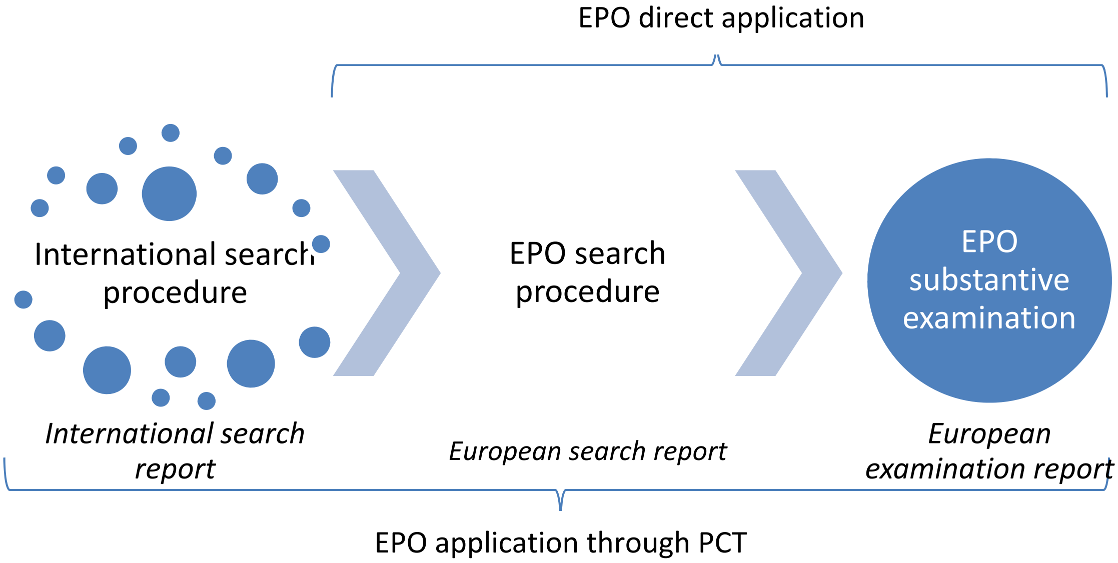


Figure 2 - Patent count in Patstat ed. October 2012

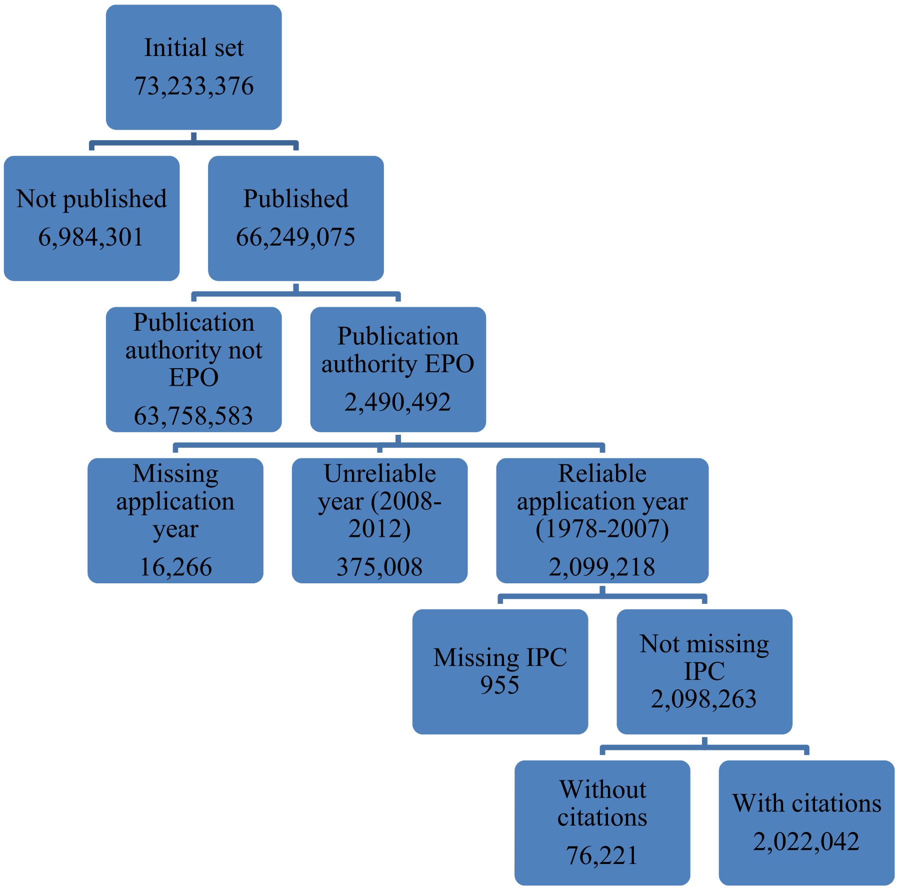


Figure 3 - Backward citations in the 2,022,042 valid patents

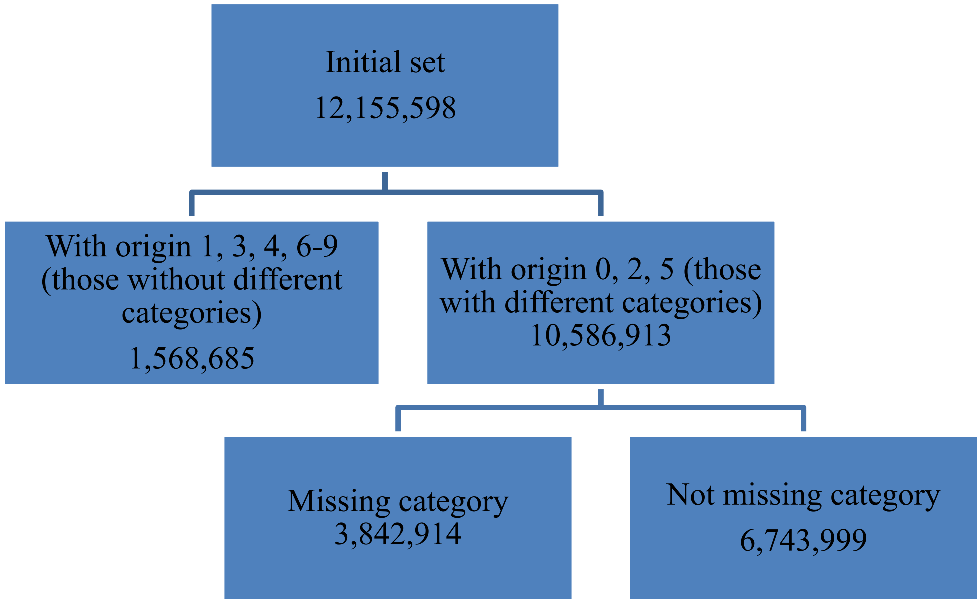


Figure 4 - Observations for the estimations

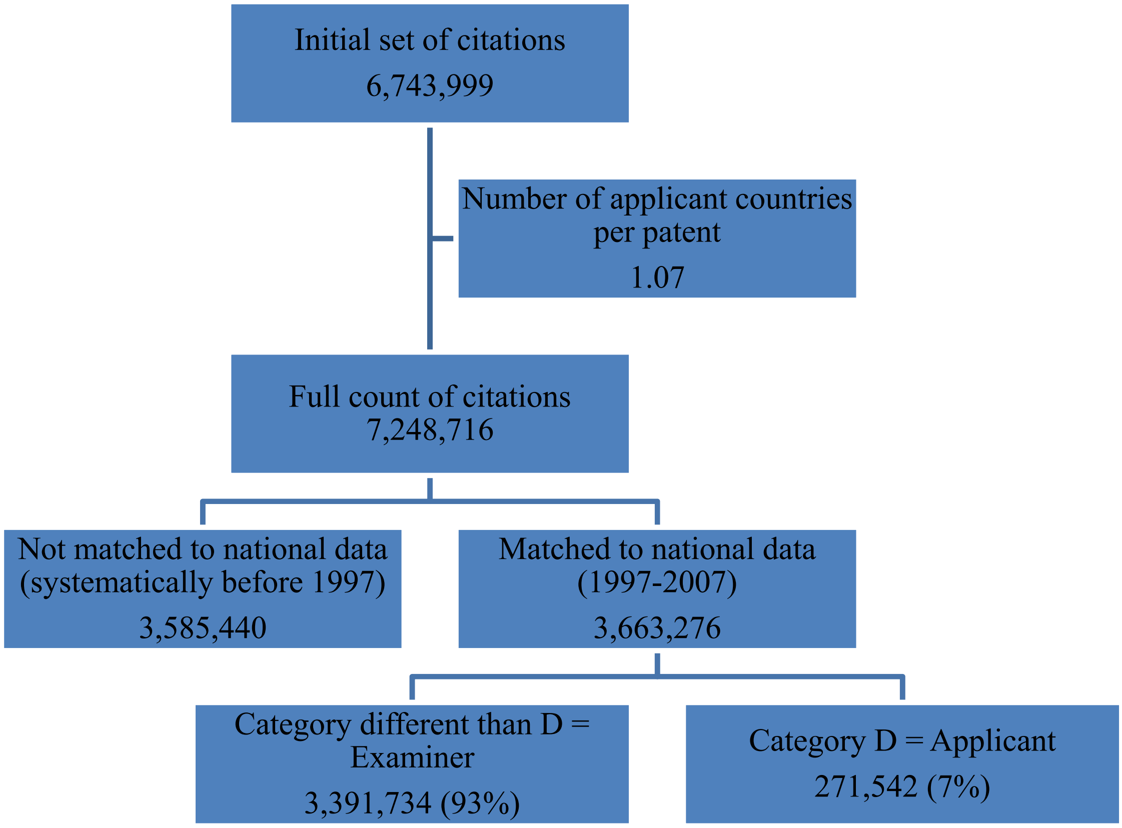
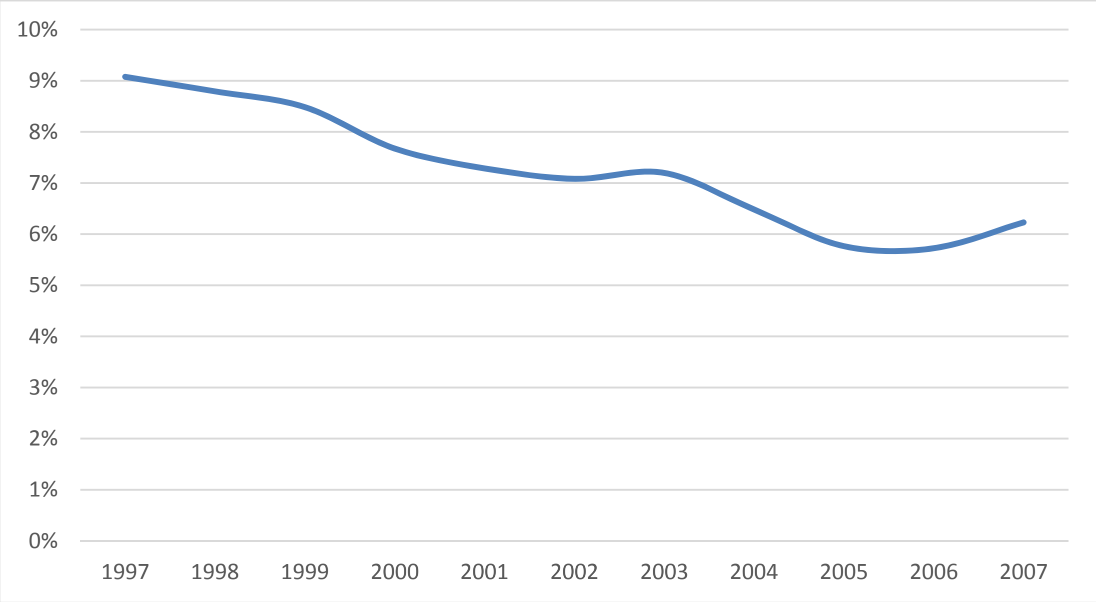
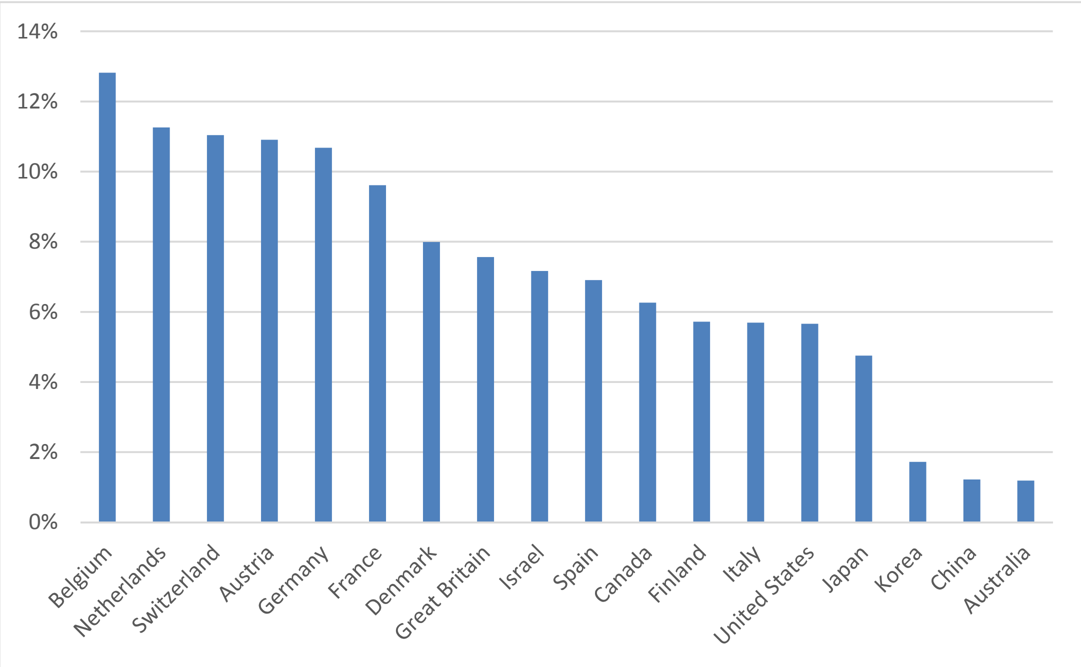


Figure 5 – Steady decline of applicant citation shares



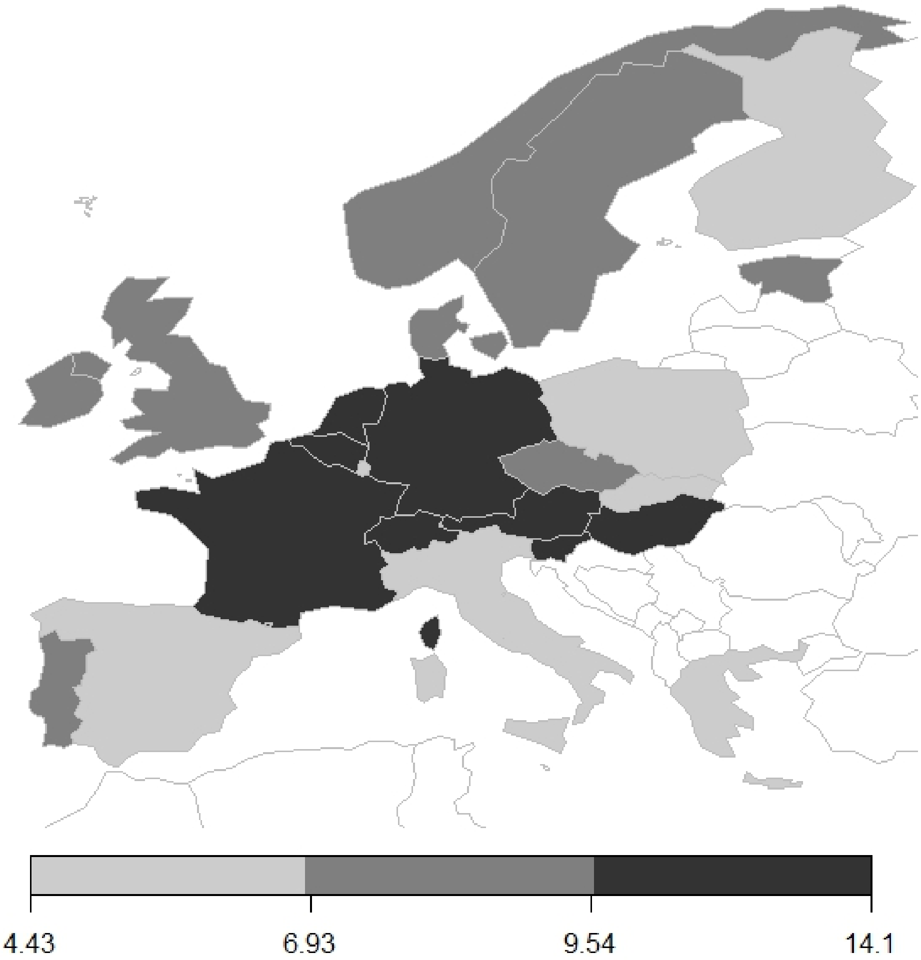
N=3,663,276. Weight: share of number of applicant countries.

Figure 6 - Large national variation in applicant citation shares



N=3,663,276. Only countries with at least 1% of EPO applications. Weight: share of number of applicant countries.

Figure 7 - The core-periphery structure of applicant citation shares in Europe



N=3,663,276. Weight: share of number of applicant countries.

Tables

Table 1 - Citation categories in an EPO search report

Category	Meaning (in search reports established for a European patent application, or in the European Patent Register)	Distribution (n=3,663,276)
A	Technological background. Used for a document representing "state of the art not [regarded as] prejudicial to the novelty or inventive step of the claimed invention."	43.31%
X	Highest possible level of relevance. "Applicable where a document is such that when taken alone, a claimed invention cannot be considered novel or cannot be considered to involve an inventive step."	32.71%
Y	Document particularly relevant if combined with another "Y" document.	12.48%
D	Document cited in the application. i.e. cited by the applicant itself.	6.54%
P	Intermediate documents, i.e. "[documents] published on dates falling between the date of filing of the application being examined and the date of priority claimed, or the earliest priority if there is more than one. Such a document may be relevant if the claimed priority is not valid.	3.23%
E	Potentially conflicting patent documents, i.e. document "bearing a filing or priority date earlier than the filing date of the application searched ... but published later than that date and the content of which would constitute prior art relevant to novelty."	1.08%
T	Documents relating to the theory or principle underlying the invention.	0.30%
I	This category is not used in search reports but in the European Patent Register for a single "X" document "particularly relevant for reasons of inventive step". See category X.	0.21%
L	Documents cited for other reasons (than the other codes). For example, if an examiner "considers that a publication, although undated, is highly relevant to the invention and can therefore be considered to be of interest to the applicant or third parties, he may choose to cite the publication in the search report as an "L" document. The search report and the written opinion should explain why this document was cited."	0.15%
O	Non-written disclosure.	0.01%

Source: column 2 (meaning), adapted from Guidelines for Examination in the European Patent Office September 2013; column 3 (distribution), own elaboration (sum equals 100%).

Table 2 - Variable definitions and descriptive statistics (n=3,663,276)

Vector	Name	Source	Variables	Description	Mean	Std. Dev.	Min	Max
$appcit_{ijklt}$	Applicant citation	Patstat	Citation category D	1 if citation category is D, 0 if other category	0.07	0.26	0.00	1.00
X_{it}	Citation characteristics	Patstat	Non-patent literature	1 if non-patent literature, 0 if patent literature	0.36	0.48	0.00	1.00
			European search report	1 if origin in search report	0.85	0.36	0.00	1.00
			Examination report	1 if origin in examination	0.00	0.06	0.00	1.00
X_{jt}	Patent characteristics	Patstat	Euro-PCT	1 if EPO-PCT, 0 if direct EPO	0.46	0.50	0.00	1.00
			Grant	1 if granted, 0 otherwise	0.18	0.39	0.00	1.00
			Filing year	Application year	2001.94	3.03	1997.00	2007.00
			A Human Necessities	1 if IPC code is A Human Necessities	0.21	0.41	0.00	1.00
			B Performing Operations; Transporting	B Performing Operations; Transporting	0.26	0.44	0.00	1.00
			C Chemistry; Metallurgy	C Chemistry; Metallurgy	0.22	0.41	0.00	1.00
			D Textiles; Paper	D Textiles; Paper	0.02	0.14	0.00	1.00
			E Fixed Constructions	E Fixed Constructions	0.04	0.19	0.00	1.00
			F Mechanical Engineering; Lighting; Heating; Weapons; Blasting	F Mechanical Engineering; Lighting; Heating; Weapons; Blasting	0.14	0.34	0.00	1.00
			G Physics	G Physics	0.26	0.44	0.00	1.00
			H Electricity	H Electricity	0.25	0.43	0.00	1.00
X_{kt}	Applicant characteristics	ECOOM*	Individual	1 if institutional sector is… Individual only	0.08	0.26	0.00	1.00
			Government	Government only	0.03	0.16	0.00	1.00
			University	University only	0.02	0.15	0.00	1.00
			Hospital	Hospital only	0.00	0.04	0.00	1.00
			Company-government	Company and government	0.00	0.04	0.00	1.00
			Company-university	Company and university	0.00	0.00	0.00	1.00
			Company-hospital	Company and hospital	0.00	0.01	0.00	1.00
			Government-university	Government and university	0.00	0.00	0.00	1.00
			# applications	Number of applications (millions)	0.00	0.00	0.00	0.15
X_{lt}	Country of applicant characteristics – economic and R&D	OECD R&D Statistics	GDP	Real Gross Domestic Product (GDP): billion Euro	0.04	0.04	0.00	0.13
			GDP per capita	GDP: Euro per inhabitant (millions)	0.03	0.01	0.00	0.07
			GERD intensity	Total intramural Gross R&D expenditure (GERD): Millions of Purchasing Power Standards (PPS) at 2000 prices	2.51	0.47	0.28	4.58
			% business funding of R&D	Business R&D funding: Share of GERD	0.64	0.09	0.17	0.91
	Country of applicant characteristics – related to EPO	EPO Annual Reports	Prob EPO exam same country	Probability of examiner from same nationality	0.10	0.10	0.00	0.26
			EPOrg member	EPO member (yes/no)	0.44	0.50	0.00	1.00

* Methodology for construction of ECOOM data explained in DuPlessis et al. (2009), Magerman et al. (2009) and Peeters et al. (2009). Non-profit organisations are subsumed within "Government". Weight: share of number of applicant countries.

Table 3 - Logistic regression of the probability of an applicant originating a citation rather than the examiner

	1 Citation and patent characteristics	2 + Applicant characteristics	3 + Country characteristics
Euro-PCT	-0.68***	-0.68***	-0.48***
	(0.01)	(0.01)	(0.01)
European search report	-0.93***	-0.94***	-0.57***
	(0.01)	(0.01)	(0.01)
Examination report	-2.73***	-2.74***	-2.43***
	(0.09)	(0.09)	(0.09)
Grant	0.30***	0.29***	0.29***
	(0.00)	(0.00)	(0.00)
Non-patent literature	0.06***	0.05***	0.10***
	(0.01)	(0.01)	(0.01)
Individual		-0.15***	-0.21***
		(0.01)	(0.01)
University		0.04***	0.08***
		(0.01)	(0.01)
Government		0.13***	0.05***
		(0.01)	(0.01)
Hospital		-0.39***	-0.31***
		(0.07)	(0.07)
Company-government		-0.09*	-0.10*
		(0.05)	(0.05)
Company-university		1.16***	1.14***
		(0.29)	(0.30)
Company-hospital		0.48*	0.31
		(0.27)	(0.27)
Government-university		-0.17	-0.40
		(0.52)	(0.53)
# applications		-0.91	-10.64***
		(0.56)	(0.57)
GDP			0.84***
			(0.14)
Per capita GDP			18.77***
			(0.86)
GERD intensity			0.27***
			(0.01)
% business funding of R&D			-0.88***
			(0.04)
Prob EPO exam same country			0.61***
			(0.04)
EPOrg member			0.64***
			(0.01)
Constant	37.53***	38.74***	67.19***
	(1.44)	(1.45)	(1.64)
Observations	3,663,276	3,663,276	3,663,276
Log likelihood	-848,023	-847,774	-838,745
χ^2	54,181	54,658	75,414
Prob>χ^2	0.000	0.000	0.000

* p<0.1; ** p<0.05; *** p<0.01. Robust standard errors in parenthesis. No collinearity according to Variance Inflation Factors. All models include a trend and eight IPC section dummies. Weight: share of number of applicant countries.