The Internationalization of Inventive Activity: A Gravity Model Using Patent Data

Picci, Lucio

European Commission - JRC IPTS, University of Bologna - Department of Economics

15 December 2008

The Internationalization of Inventive Activity: A Gravity Model using Patent Data

Lucio Picci[1]

This version: 1 December 2008

Abstract

This paper discusses the extent and the determinants of the internationalization of European inventive activity, between 1990 and 2004, using an innovative method to treat the information contained in the European Patent Office's Patstat database.

The observed level of internationalization of inventive activities, while being rather low, has steadily increased over time. The amount of collaboration between actors residing in different countries is assessed by means of a "gravity model", as it is familiar in the literature on international trade. The amount of bilateral collaboration is positively affected by the presence of a common language and a common border, and by the common participation in the European Union. Participation in the Euro Zone is also found to have a (marginally) negative effect.

International collaboration is negatively affected by distance, with estimated elasticities that are significantly smaller than the ones that characterize international trade. Contrary to the rumors about the "death of distance", this effect has become stronger in recent years.

---

[1] European Commission - DG JRC, Institute for Prospective Technological Studies, C/ Inca Garcilaso, s/n - E-41092 Sevilla, and Department of Economics, University of Bologna, Strada Maggiore 45, 40125 Bologna. E-mail: lucio.picci@ec.europa.eu. I thank Marc Bogdanowicz, Albert Hu, Andrea De Panizza and Antoine Schoen for useful discussions on a previous version of the paper. The opinions expressed do not necessarily coincide with those of the European Commission.

1. Introduction

Almost all contemporary products are the result of some form of international collaboration and trade. A cell phone is a bundle of many components that, almost invariably, are produced in different countries. Raw materials are traded to such an extent that anything containing plastic or metal components is also very likely to be the result of international trade. As for services, the current lore on international outsourcing of back-office routines indicates that internationalization, while not being preponderant, is certainly increasing.

Innovative activities, on the other hand, are largely the result of efforts which take place in single countries. Almost twenty years ago, considering the patenting activities of a sample of big US firms, Patel and Pavitt (1991) observed that "in most cases, the(ir) technological activities […] are concentrated in their home country", and concluded that "despite being a critical resource in the global competition and performance of both companies and countries […] the production of technology remains far from globalised." Since then globalization has certainly made inroads into this domain, thanks to robust increases in international collaboration of various types, to the off-shoring of many R&D labs, and to advances in the protection and trade of intellectual property rights. However, as we will demonstrate, it is still true that innovative activities are predominantly national in scope.

This paper aims to identify the determinants of the intensity of international collaboration in inventive activities, and what they tell us about why the degree of internationalization of innovative activities is still relatively low. To address this issue, several researchers have exploited in various ways the information contained in patent data (see, among others, Patel and Pavitt, 1991; Patel and Vega, 1991, and Le Bas and Serra, 2002). The research presented here also uses patents data. While most previous studies have considered the patent portfolios of (multinational) firms, here we attribute patents to countries, by exploiting the fact that patents data provide separate information on the nationality of the inventors and of the applicants. If a given patent involves only people and organizations residing in the same country, we define it as "national". If, on the other hand, at least one inventor or one applicant resides in a country different from that of the others, then we call the patent "international". Moreover, the presence of an important functional distinction between inventors and applicants allows us to define and analyze different measures of internationalization.

Patent data, as a measure of inventive output, have virtues and shortcomings (Smith, 2005, and Griliches, 1990). Many innovations, particularly of production processes, do not result in any patent applications, and firms often prefer to protect their inventions by keeping them secret, rather than by asking for the protection afforded by patents. However, this limitation of patent statistics is less important when the focus is on international innovation, because in that case the propensity to patent is bound to be higher, given that trade secrets are more difficult to keep in situations where the innovators reside in several countries and may belong to distinct organizations.

While not all innovations are patented, the opposite is also true - that is, not all patented inventions produce innovations. Patents may have very different values, and for each superstar patent, which introduces a very relevant and successful product or process, there are countless others with limited or no use. Possibly a more serious problem is the fact that patenting activity is increasingly just one of many strategies that oligopolistic firms have at their disposal. "Defensive patenting", in particular, has become a way of accumulating a sizeable patent portfolio to be used as a bargaining chip. Indeed, at times the collections of patents owned are better summarized as ammunition for lawyers, than blueprints for engineers. While the worst excesses of such “patent inflation” are probably confined to the United States, the issue is obviously a very serious one (see Jaffe and Lerner, 2004 for an assessment and a critique; and Archontopoulos et al, 2007 for a quantitative assessment of the problem at the European Patent Office (EPO)).

One way of filtering out low-quality patents is to only consider the applications presented at particularly important patent offices. Another possibility is to consider "triadic patents", meaning all patents filed at least at the EPO, the United States Patent and Trademark Office and the Japan Patent Office (see Dernis and Khan, 2004). Since a triple filing is quite expensive, it may be expected that applicants choose to incur the related costs only when they believe that their invention deserves it. However, this approach does not filter out strategic patenting activities, and in fact may achieve the opposite, since such strategy is more likely to address the patenting offices covering the widest markets.

Approaches that aim to exclude low-quality patents also overlook relevant inventive activities. Even when they do not represent a genuine advancement of the technological frontier, patents hint at the presence of an absorptive capability, because learning about existing technologies is an integral part of R&D activities. As Cohen and Levinthal (1989) put it, R&D is always “innovation and learning”. For this reason, we have adopted a methodology that computes patent statistics inclusively, rather than selectively, by focusing on the (priority) applications presented at any one of the EU27 national patent offices and at the European Patent Office (EPO).

We have adopted a gravity model to study the determinants of the intensity of collaboration between pairs of countries. The gravity model ("one of the more successful empirical models in economics", Frankel and Rose, 2002) has had remarkable success in explaining bilateral trade flows. It describes trade between two countries as increasing in their economic size, and as decreasing in their distance (hence the resemblance with Newton's gravity law). This model has been given several theoretical rationalizations, the first proposed by Anderson (1979). Disdier and Head (2008) performed a meta-analysis on 103 papers applying the gravity model and report a mean elasticity of bilateral trade with respect to distance of 0.9, "indicating that, on average, bilateral trade is nearly inversely proportionate to distance". They also show that distance effects "decreased slightly between 1870 and 1950 and then began to rise". Besides the negative effect of distance, other variables have generally been found to positively determine bilateral trade, such as the presence of a common border, a common language, a common market, and a common currency area (for the latter, see Rose, 2000, and Baldwin, 2006). The gravity model has been also applied to describe relations other than traditional trade in goods. In particular, distance effects have been found in the trade of services (Ceglowski, 2006, and Kimura and Lee, 2006), trade through the Internet (Berthelon and Freund, 2008), knowledge flows through patent citations (Peri, 2005), immigration flows (for a recent example, see Lewer and Van den Berg, 2007), and also in fields outside economics.

Our approach appears to have only one partial antecedent. This is described in a paper by Guellec and van Pottelsberghe de la Potterie (2001), whose measures of R&D internationalization present similarities with ours, to be illustrated in the next section. However, apart from the fact that today almost a decade more of data is available for analysis, our research differs in many respects from the previous literature. One dimension in which this paper is innovative is in its treatment of patent information. We dedicate Section 3 to a description of these important methodological aspects. In Section 4, we illustrate the empirical model and its econometric estimates. Section 5 discusses the results, and Section 6 offers our conclusions.

2. Measures of internationalization

To present our measures of internationalization in R&D, we first briefly review the familiar concept of fractional counting of patents (see for example Dernis et al., 2001). To help make the discussion as easy to follow as possible, we use a simple fictitious example. We consider three countries, United Stated (US), France (FR), and Germany (DE), that in a given year produce a total of P=3 patents. Column I in Table 1 indicates the nationality of the inventors and applicants that contributed to these three inventions.

Table 1 about here

In order to assign patents to countries, two alternative criteria may be chosen: either according to the nationality of the applicant(s), or of the inventor(s). The former defines the "applicant criterion" and the latter the "inventor criterion". Whenever an application has more than one inventor or applicant, some of them coming from different countries, patent assignment is carried out by resorting to fractional counts. So, for example, patent n. 1 counts as ½ German and ½ American according to the applicant criterion, and ½ American, ¼ German and ¼ French according to the inventor criterion.

Let us call $Inv_{i,p}$ the fraction of patent p attributed to country i according to the inventor criterion, and $App_{i,p}$ the analogous measure according to the applicant criterion.[2] Column II and III of Table 1 report these measures for the three patents. For each patent application, the sum of all the country's contribution according to the inventor criterion has to be equal to 1: for each patent, $Inv_{US,p} + Inv_{DE,p} + Inv_{FR,p} = 1$, where the first subscript indicates the country, and the second the patent. These sums are indicated in Column IV of Table 1.

The total fractional assignment of the three patents to each country is simply equal to the sum of the individual assignments:

(1) $$Inv_i = \sum_{p=1}^{P} Inv_{ip}$$

and:

[2] When considering the fictitious example, instead of the subscript i we will use the mnemonic symbol of the relevant country. Also, for clarity we omit in all cases a time subscript, that should always be present.

(1') $App_i = \sum_{p=1}^{P} App_{ip}$

They are reported in the last two rows of Table 1. For example, Germany produced a total of 0.75 patents according to the inventor criterion, and of 0.5 patents according to the applicant criterion.

Having defined the concept of fractional counting, we proceed to develop useful measures to express the degree of collaboration in producing patents between (the agents residing in) two countries. Again, we discuss them using the fictitious example of Table 1. We define three concepts of internationalization of a given patent:

Inventor internationalization (in short, InvInv): Whenever a patent has at least two inventors resident in different countries. All three patents in our example display this type of internationalization.

Applicant internationalization (AppApp): Whenever a patent has at least two applicants resident in different countries. Patent 1 and 2, but not patent 3, display this type of internationalization.

Inventor-Applicant internationalization (InvApp): Whenever a patent has (at least) one inventor and one applicant that come from different countries. All three patents of the example belong to this type.

In a patent, the presence of InvInv implies InvApp – if there are two inventors from different countries, then it is also necessarily true that one inventor and one applicant are from different countries. Similarly, AppApp implies InvApp. The relation between the three types of internationalization is depicted in Figure 1. It indicates, in particular, that all international patents are necessarily of the InvApp type, and possibly also of the InvInv and/or of the AppApp type.

Figure 1 about here

An analysis of any shortcomings of our concepts of internationalization should be carried out with an eye to the alternatives available. As we mentioned in the introduction, there are two competing approaches to analyzing internationalization of R&D activities using patents data. One is by assembling a firm's portfolio: Firms are typically selected (also) according to their size, and this leads to problems of sample selection. We, on the other hand, look at patents regardless of the size or type of the applicant(s), and resort to an "automatized" criterion to select international patents. The limits each approach may have ultimately derive from the fact that patent applications are so numerous and are not amenable to a case-by-case examination.

There are two forms of international inventive effort that our approach may fail to detect. First, imagine that a firm owns an R&D unit in a foreign country, producing an invention with the help of inventors that are all resident in that same location. If, moreover, the applicant of the filing is the foreign subsidiary (instead of the firm's headquarters), or a subsidiary located in the home country, then all the applicants and the inventors would be from the same country and therefore the patent application, according to our taxonomy, would fall into the "national" category. However, note that usually, as we will verify in the next section, multinational firms apply for their patents through their headquarters – thus, the patent in this example would fall into the InvApp type. Another case of internationalization that would go undetected is when two firms from different countries constitute a joint R&D effort in one of the two countries, or in a third country, and produce an invention where all the inventors are residents of the country where the jointly-owned firm is registered. Arguably, there should not be very many of these cases. Moreover, it is possible that researchers from both countries would team up in the jointly-owned entity, so that their patenting activities would show up as inventor and inventor-applicant internationalization. Also, there may be patents that we classify as international, which, in fact, are not. For example, a multinational corporation (MNE) could have its legal headquarters in one country, but most of its operations in another. In this case, its patents would automatically display applicant internationalization. In the next section, a careful analysis of a sample of international patents will lead us to conclude that, overall, the number of problematic cases should be quite limited.

Having discussed the general concepts of Inventor, Applicant, and Inventor-Applicant internationalization, we now proceed to define the related measures, starting with the first one. For each patent, the strength of the relation between inventors in country *i* and *j* is expressed as the product of the attribution of that patent to the two countries:

(2) $$InvInv_{ijp} = Inv_{ip} \cdot Inv_{jp}$$

This measure attributes a greater weight to collaborations where the two countries have more similar weights. So, for example, the collaboration between the US and France is equal to ½ · ¼ = $^1/_8$ in patent n. 1 (where there are 1 French and 2 American inventors) and to ½ · ½ in patent 3 (where the total number of inventors, 2, is equally divided between the US and France. In fact, if *i* is different from *j*, $0 \leq InvInv_{ijp} \leq 1/4$, where the upper bound is reached when the total number of inventors is equally divided between two countries, and the lower limit applies when a patent is national.

The aggregate strength of the relation between the inventors of two countries is defined as the sum of the above, over all patents:

3) $$InvInv_{ij} = \sum_{p=1}^{P} Inv_{ijp}$$

Below, we report the values for all the combinations of the three patents in Table 1, where for clarity, instead of the indexes *i* and *j*, the acronyms of the countries are employed.

$$InvInv_{US,US} = 0.5 \cdot 0.5 + 0 \cdot 0 + 0.5 \cdot 0.5 = 0.5$$

$$InvInv_{US,DE} = 0.25 \cdot 0.5 + 0 \cdot 0 + 0 \cdot 0 = 0.125$$

$$InvInv_{US,FR} = 0.25 \cdot 0.5 + 0 \cdot 0 + 0.5 \cdot 0.5 = 0.375$$

The top part of Table 2 shows the values of these interactions for all three cases. Note that $Inv_{ij} = Inv_{ji}$ (the order of the countries is irrelevant). Using (1), it is easy to see that:

(4) $$\sum_{i=1}^{N} InvInv_{ij} = Inv_j \quad \text{and} \quad \sum_{j=1}^{N} InvInv_{ij} = Inv_i$$

For example, as predicted by (4):

$$InvInv_{US,US} + InvInv_{US,DE} + InvInv_{US,FR} = 0.5 + 0.125 + 0.375 = 1 = Inv_{US}$$

These sums are reported for all three countries in the last column and in the last rows of the top part of Table 2, and correspond to the values reported in Table 1. They show that the country patent portfolio, assigned according to the inventor criterion, may be expressed as a sum of pairwise measures of country inventive collaboration ($InvInv_{ij}$).

The measure of applicant internationalization is constructed along the same lines, and the following formulae hold:

(2') $$App_{ijp} = App_{ip} \cdot App_{jp}$$

(3') $$App_{ij} = \sum_{p=1}^{P} App_{ijp}$$

(4') $$\sum_{j=1}^{N} App_{ij} = App_i \text{ and } \sum_{i=1}^{N} App_{ij} = App_j$$

All computations for this case are shown in the middle part of Table 2. Note that $App_{ij} = App_{ji}$ (again, the order of the countries is irrelevant). Equation (4') allows us to express a country patent portfolio, according to the applicant criterion, as a sum of interactions between applicants in different countries. The values reported in the last column and row of the middle part of Table 2 correspond to those of Table 1.

We construct a measure of Inventor-Applicant internationalization similarly. The strength of the collaboration between inventors in country *i* and applicants in country *j*, for a single patent *p*, is defined as:

(5) $$Invapp_{ijp} = Inv_{ijp} \cdot App_{ijp}$$

Summing over patents provides a measure of the strength of the overall collaboration between country *i* inventors and country *j* applicants:

(6) $$Invapp_{ij} = \sum_{p=1}^{P} Inv_{ijp} \cdot App_{ijp}$$

These measures aggregate to the patent attributed to a country either according to the inventor, or to the applicant criterion, depending on whether the summation is over *i*, or over *j*:

(7) $$\sum_{j=1}^{N} Invapp_{ij} = Inv_i$$

(7') $$\sum_{i=1}^{N} Invapp_{ij} = App_j$$

The bottom part of Table 2 indicates all computations for our fictitious example. Note that $InvApp_{ij}$ generally differs from $InvApp_{ji}$.

The quantities defined in (3), (3') and (6) are the three measures of internationalization of innovative activities that we will assess in Section 4 using a gravity model. In the next section, in order to provide a first description of the degree of internationalization, we will use *relative* measures of internationalization, expressed as a share of the total number of patents. It is straightforward to construct relative measures of (3) and (3'):

(8) $Inv_{ij|i} = Inv_{ij} / Inv_i$

and

(8') $App_{ij|i} = App_{ij} / App_i$

where $\sum_{j=1}^{N} Inv_{ij|i} = 1$ and $\sum_{j=1}^{N} App_{ij|i} = 1$.

There are in fact two conditional measures of inventor-applicant internationalization, depending on whether the normalization is carried out with respect to the inventors of country *i*, or to the applicants of country *j*:

(9) $Invapp_{ij|i} = Invapp_{ij} / Inv_i$

(9') $Invapp_{ij|j} = Invapp_{ij} / App_j$

where $\sum_{j=1}^{N} Invapp_{ij|i} = 1$ and $\sum_{i=1}^{N} Invapp_{ij|j} = 1$.

Our metrics of relative internationalization have similarities with those of Guellec and van Pottelsberghe de la Potterie (2001), who adopt three measures that they call SHAI, SHIA, and SHII. The first one is similar to our $Invapp_{ij|i} = Invapp_{ij} / Inv_i$, the second to $Invapp_{ij|j} = Invapp_{ij} / App_j$, and the third to $Inv_{ij|i} = Inv_{ij} / Inv_i$. Our $App_{ij|i} = App_{ij} / App_i$ has no analogue in their paper. There are, however, several differences in the way that the measures are constructed, perhaps the main one being that, here, fractional counts of patents lead to counting as "more international" those patents where international collaboration is more pronounced. One advantage of our measures is that they are coherent with the concept of fractional counting, in that they allow us to express country patent counts as sums of pairwise internationalization linkages (equations 4, 4', 7 and 7'). The measures adopted by Guellec and van Pottelsberghe de la Potterie (2001), on the other hand, do not make this distinction, and consider alike all patents where there is at least some international collaboration of a given type. Similar considerations hold for the patent statistics of internationalization presented in OECD (2008).

3. The data

We consider patent applications, not granted patents. This choice, besides being a common practice in the literature, has the advantage of allowing an analysis of more recent data, considering that several years typically elapse between the filing and the granting of a patent. The source of the data is the Patstat database, based on the EPO's "master bibliographic database DocDB" (European Patent Office, 2008a and 2008b). It covers data from more than 80 patent offices and contains a wealth of information on individual filings, including the identity and country of residence of applicants and inventors (European Patent Office, 2008b). An important and innovative characteristic of Patstat is that it allows the identification of patent applications that claim the right to priority (that is, it permits us to distinguish between multiple applications for the same inventions in several patent offices). We take full advantage of this characteristic and we only consider "priority" applications. Excluding all other applications allows us to assess only those filings that reflect true inventive activity that is centered in the geographic area of interest – in our case, the European Union.

Most studies on patents only use information from a single patent office. Given our focus on the internationalization of patents, this approach would not be viable, because the presence of the "home bias effect" would vitiate any study based on a single, even if very important, national patent office. For this reason, we consider applications filed in any one of the patent offices of a member state of the European Union (in its current configuration of 27 States, EU27), or at the EPO. Guellec and van Pottelsberghe de la Potterie (2001) and OECD (2008) focus their attention on the filings to the EPO only, a strategy that rests on its continental scope. However, such a choice also is questionable, for two reasons. First, many European patents never reach the EPO as priority filings, or as successive applications. Secondly, the decision to file to the EPO is influenced by factors that are country-dependent. In particular, agents from smaller countries, which have been members of the EPO for longer and whose national patent offices have higher filing fees, have a higher propensity to use that institution (De Rassenfosse and van Pottelsberghe de la Potterie, 2007).[3] The two reasons together imply that only focusing on the EPO may induce biased estimates.

We would like to obtain a better understanding of the nature of the patent applications that we characterize as "international". Their sheer quantity (a total of 72,291 patents) makes it impossible to carry out a case-by-case analysis. As a compromise between what would be desirable, and what can reasonably be achieved, we selected, out of the population of 72,291 international patent applications, a stratified random sample of 300 patents, with proportional allocation in the year variable. For each one of them, we determined the type of applicant(s), according to the following classification.

"Multinational Enterprises" (MNEs) are firms that control at least one production unit in a country different from where they are based. A distinction is made between headquarters (HC) and subsidiaries (SUB). Also, subsidiaries that are registered in the same countries as the headquarters are recorded separately. "Firms" are all firms that are not (part of) a MNE. "Universities" include all universities, and "Public research institutes" are research institutions that are primarily financed by public administrations (in all cases considered, they are also publicly owned, and clearly public in nature).[4]

Table 3 about here

[3] Using data from the EPO only (or from one or few national patent offices, as in the triadic approach) was also a matter of convenience and, until some time ago, it was almost unavoidable. With the maturity of the Patstat database, better alternatives have become available.

[4] The necessary information have been collected mainly through the Internet, taking advantage of the fact that most sites of MNEs report fairly accurate details of their organizational structure and of their history. We also used the Amadeus business registry to solve a few difficult cases. In a total of 7 cases, most of them corresponding to the first years of the period under study, there remained some doubts as to whether a firm was a MNE or not. In all of them, the firm was declared not to be a MNE, what seemed to be the more likely case.

Table 3 reports the relative frequencies (in percentage terms) of the type of applicant conditional on the type of internationalization (note that the sum of the percentages by columns is always equal to 100%). At the bottom of the Table we also report the relative share of the types of internationalization, both in the sample and in the population. Around 90% of the patents that are only of the InvApp type (and not of the InvInv and/or AppApp type) are the result of the activity of MNE. In most cases, MNE file through their headquarters, and only in 4.57% of the cases (corresponding to 5.2% of the "InvApp only" applications filed by MNE), the applicant is a MNE subsidiary from a country different from the home country. One implication of the fact that MNEs mostly file through their headquarters is that the number of international patents that escape our definition of internationalization because they are the result of the filing of a MNE subsidiary abroad, and of the work of inventors that are all from the same foreign country – remember our discussion in the previous section – is likely to be negligible. Note that none of the 300 applications, that are of the InvApp type only, was filed by a person, by a public research institute, or by a university.

The next three columns of Table 3 show the share of the different types of applicants for international patents that, besides being of the InvApp type, also present InvInv or AppApp internationalization, or both. From the percentages reported at the bottom of the Table (both for the sample, and for the whole population of international patents), it can be seen that there are only few occurrences of applicant internationalization (less than 9% of international patents, in the population). For international patents that also display inventor and, in particular, applicant internationalization, the relative share of MNEs is smaller. We also observe that the only filings by universities and by public research institutes in the sample always involve the presence of inventors that are from different countries and, in a very few cases, also the collaboration of applicants from different countries. A Pearson χ^2 test shows that the relative importance of each type of applicant is not independent to the type of internationalization, implying that the differences that we have commented upon are statistically significant overall.

Next we consider whether there is a relation between the type of applicant and their filing strategies. We distinguish between three types of patents in terms of the filing strategy of their applicants: 1) patent applications that are never filed to the EPO; 2) patent applications that receive their priority filing at the EPO; and 3) patent applications that, while receiving a priority at any of the EU27 national patent offices, are eventually mentioned as prior applications that are also filed at the EPO. Table 4 shows the results. The bottom of the Table indicates that about one third of all international patents never reach the EPO. Indeed, as we discussed in the previous section, only focusing on the EPO excludes much relevant inventive activity.

Table 4 about here

We find that MNEs have a higher propensity than other firms to first file to the EPO. On the other hand, in our sample there is no university, research institute or person that applied first to the EPO. The relation is marginally insignificant, as the first of the two Pearson χ^2 tests indicates. We observe that types of organizations that less frequently do their first filing to the EPO often do so later on, by means of successive filings of the same, or of related, inventions. This is shown in the far right column of Table 4, where we still observe a slightly higher share of MNEs. As the second Pearson χ^2 test indicates, however, this difference is not statistically significant.[5]

We now focus our attention again on the whole population of patents, in order to assess the overall degree of internationalization of European inventive activity between 1990 and 2004. To this end, we compute the relative measures of internationalization defined by Equations 8, 8', 9 and 9' of the previous section, aggregating the 27 countries that constitute the European Union. The results are shown in Figure 2.

Figure 2 about here

---

[5] The stratified structure of the sample also allows us to consider any changes in time in the relations that we have discussed. Dividing the time period under consideration into two intervals (1990-1997 and 1998-2004), we find that: a) the propensity to first file international patents to the EPO, instead than to a EU27 national patent office, is significantly greater in more recent years b) however, once successive filings are taken into consideration, while the fraction of application that eventually reaches the EPO is still higher in more recent years, a Pearson χ^2 test statistics does not reach conventional significance levels; c) we do not find any evidence that the composition of the type of applicant changes significantly between the first and the second part of the time period considered.

All measures indicate a degree of internationalization that is still rather modest, even though it has increased steadily since the early 1990s. Also, we observe important differences among the four alternative metrics, with the two relative measures of inventor-applicant internationalization being well above the others. The first of the two (indicated in the Figure as InvApp | Inv) refers to patents that have European inventors and extra-European applicants, and AppInv | App is the opposite. The typical case accounted for by the InvApp | Inv measure is the one of the extra-European MNE owning an R&D lab in Europe and filing the patents produced there through its headquarters in the home country. The fact that the first measure is higher than the second indicates the relatively high importance of extra-European applicants in European inventive activity. Applicant internationalization is very low, always below 1%, indicating that only in a very small minority of cases applicants from different countries jointly file an application. The measure of inventor internationalization at the end of the period is slightly above the 3% mark.

Overall, these data allow us to declare that the degree of "globalization" in the production of technology, to refer again to the work of Patel and Pavitt (1991), has certainly increased since the time when they wrote, but is still rather low in relative terms.[6]

4. A gravity model of international inventive activity

The basic model that we estimate is the following:

$$\ln(INT_{ijt}) = \beta_0 + \beta_1 \ln(A_{it}) + \beta_2 \ln(A_{jt}) + \beta_3 \ln(dist) + \lambda L_{ijt} + \beta_4 D_i + \beta_5 D_j + \beta_6 DT_t + \varepsilon_{ijt}$$

[6] While in the previous Section we clarified that there are important differences between our measures of internationalization, and the ones in Guellec and van Pottelsberghe de la Potterie (2001), still it is useful to provide an example of how they compare to one another. When looking at the EPO, they find, for Germany and for the years 1993-1995, values of their measures SHAI, SHIA and SHII of respectively 6.3%, 4.5% and 6.6%. In the present case, the more directly comparable measure are equal, for the year 1994, to 2.0%, 1.9% and 0.6%. The difference is probably attributable mainly to the fact that the EPO afforded, particularly in those early years, a very biased sample of European inventive activity, as it was discussed in the previous section.

where *ln* is the natural log, INT_{ijt} is one of the three bilateral measures of internationalization of inventive activity, A is the "inventive mass" of each country, to be proxied with appropriate patent statistics, *dist* is the distance between the capital cities of pairs of countries (computed with the great circle formula) and L is a vector of other conditioning variables. It includes dummies for the presence of common borders and of a common language[7] (with no variation in time), and for the inclusion in the European Union and in the European Monetary Union (with variations in time for several of the countries considered). We include two more time-varying variables in *L*. One is a multiplicative term of logged distance and of a time trend, to assess whether any time distance effect varies in time. The other is an indicator of pairwise "inventive proximity". It is computed as the correlation between two vectors, one for each country, formed by the number of patent priority applications in each of the eight IPC technology classes.[8] The two D variables are country-specific fixed effects, one for each country, and the variable DT is meant to capture any time shock affecting all bilateral relations.[9] This specification is quite flexible, including all possible fixed effects, short of estimating a Fixed-Effects panel model.[10] Flexibility is also added by pooling observations from different years, instead of summing them over periods of a few years, as is sometimes done in the literature. The presence of a

7 The language dummy takes fractional values for multilingual countries. For example, it is equal to one half between Belgium and France (the presence of a small German speaking minority in Belgium is ignored), and to one third for the pairs of Switzerland with Germany, France and Italy.

8 These are the eight top "sections" (A through H) of the International Patent Classification (IPC) taxonomy. See http://www.wipo.int/classifications/ipc/en/ (last visited on 1 December 2008). The correlations are computed for each year on the same dataset used for all the exercises carried out in the paper. Fractional counting is adopted to address all the cases when more than one IPC category is assigned to a patent.

9 Guellec and van Pottelsberghe de la Potterie (2001) explain bilateral relations using a model that, among other differences with respect to ours, does not include distance among the regressors (so that it is not a “gravity model”).

10 A Fixed Effect model, in its Least Square Dummy Variable representation, would simply obtain by also including dummies indexed *i,j*, that is, one for each statistical unit of the model. In this case, the parameters would be identified uniquely by the time variation of the data. All time-invariant variables (such as distance, and most likely candidates for inclusion into the L vector) would be swept away by the fixed effects, so that the model would be of very little empirical interest.

year dummy variable controls for the presence of time shocks affecting all pairwise relations in the same way.

The model is first estimated using OLS, both incorporating the whole set of pairwise measures of internationalization, and also limited to those that include only the countries that are currently members of the European Union. This is meant to address the asymmetric treatment among European and non-European countries, the latter being affected by a home bias effect because we only consider European patent offices.[11] In most cases, however, both samples provide qualitatively similar results. One problem in the estimation of gravity models derives from the presence of cases when pairs of countries do not entertain any relation – be it of trade or, as in our case, of inventive activity. Given that the model is expressed in logs, all these occurrences generate missing values, and the ensuing truncation of the dependent variable could be a source of biased estimates. To address this problem, we also adopt the two-step procedure introduced by Heckman, popularly known as the Heckit estimator (Heckman 1979). When interpreting the results, in the few cases when the OLS and the Heckit estimators suggest conflicting conclusions, the Heckit estimates will be privileged.

Table 5 about here

Table 5 shows the estimates of the gravity model on $InvApp_{ij.}$[12] Distance is found to negatively affect international collaboration. In all cases, we also find that this effect

---

[11] In principle, other patent offices could be included in the construction of the data, up to the whole set currently covered by Patstat. However, the degree of complexity of the computations as they are is already quite high. To produce the results of this paper, starting from the output of 60 SQL queries (four for each year considered, producing a total of more than 13 million records), a set of 19 Stata ".do files" run sequentially, with a total execution time of approximately 30 hours using a small workstation. Secondly, the use of Patstat still requires much care, because at times the data present idiosyncrasies that require careful examination. A line had to be drawn somewhere.

[12] In this and in all following cases, the first step of the Heckit estimator, a Probit describing the selection into the sample, used the same list of dependent variables of the gravity model, excluding the measures of inventive activity and the interaction between logged distance and time, plus GDP based on purchasing power parity, expressed at current dollars, taken from IMF's World Economic Outlook Database (http://www.imf.org/external/ns/cs.aspx?id=28, last accessed on 1 December 2008).

increased over time, as is indicated by the negative and significant impact of the cross-product of distance with the time trend. The last line of Table 5 shows, for the Heckit estimates only, the implied elasticities of internationalization with respect to distance, for the first and for the last year of the period under consideration. They range from -47% to -60% for the whole sample of countries, and from -32% to -82% for the sample of European countries.

Having a common border, sharing a common language and having more similar technologies has a significantly positive effect on international inventive collaboration, as does being a member of the European Union, when the analysis is constrained to the EU27 sample. It is particularly interesting to observe that the impact of the Euro Zone variable is estimated to be negative in the Heckit estimates, and significant in the full sample case, and not far from it in the reduced sample.

Table 6 shows the estimates of the gravity model on the $InvInv_{ij}$ measure of internationalization.[13]

Table 6 about here

We highlight the most relevant differences with respect to the previous estimates, most of which are similar. The elasticity of distance is now estimated to be more modest, up to -46%, in the last year and for the whole sample. The negative effect of distance significantly increases in time in both samples, with magnitudes that are not too different from those that we found in the previous case. Having a common border has a more pronounced effect than before. Being part of the Euro Zone now has a significantly positive effect in the reduced sample, and an insignificant effect in the broader sample.

---

13 Here the number of non-zero observations (equal to 4276 and 2726 respectively for the whole sample and for the EU27 selection) is greater than in the previous case (3958 and 2448, respectively). This does not contradict our assertion that the InvApp measure is the more "general" of the three (see also Figure 1). In fact, as the fictitious example of Table 1 also shows, $InvApp_{ij}$ may be smaller than the corresponding $InvInv_{ij}$ (or $AppApp_{ij}$) and, in particular, $InvApp_{ij}$ may be equal to zero when the other two measures are not.

Finally, we present the estimates of the gravity model on $AppApp_{ij}$. This measure involves a relatively modest number of cases, being defined by rare collaborations between applicants residing in different countries.[14] The results, shown in Table 7, indicate that the negative role of distance, while still being significant, is weaker compared to the previous cases. For the whole sample, we find a negative effect of distance, but we do not find evidence that it increased over time (to the opposite, the estimated coefficient of the product of logged distance with time is *positive*, even if of modest magnitude and insignificant). For the reduced sample, we find that the elasticity of distance increases from approximately zero at the beginning of the period, to about -43% at the end.[15]

Table 7 about here

A smaller role for distance is also indicated by the fact that the estimated coefficient on the border dummy, while being negative and significant, is of smaller magnitude than in the two previous cases. The estimated effect of sharing the same language is positive and significant. The effect of belonging to the EU is positive and significant for the smaller sample, while the effect of belonging to the Euro Zone is estimated to be positive in both cases.

5. Discussion

From our results, it emerges unambiguously that distance negatively affects the internationalization of R&D activities, with estimated elasticities that, at their highest, are about half of what is usually found in the literature on international trade. It is quite likely that the costs of travel alone are not able to explain this effect – just as they do not explain the negative effect of distance on international trade. It could well be that what really

---

[14] This also may explain why in this case the differences between the OLS and the Heckit estimates are more important than before: the lower is the ratio of non-zero observations, the more likely it is that OLS estimates will be prone to a sample selection bias.

[15] Please note that while the estimated coefficient of the log of distance is insignificant, the logged distance and the cross product of logged distance with time are jointly significant, with a p-value smaller that 1%.

matters is what may be seen as a measure of loosely defined "cultural" distance. The success of R&D activities rests upon the smooth functioning of many tasks that are not limited to research activities proper, but also include, for example, the building of the necessary infrastructure and the protection of the intellectual assets that are eventually produced. Since the necessary mutual understanding among all parties involved may be guaranteed only in part through contractual arrangements, tacit agreements end up playing an important role. Arguably, the further away two cultures are, the more difficult it is for tacit agreements to be honored, and the more problematic is collaboration.[16]

The physical distance variable could then proxy for this "cultural distance" element which is missing from our model. Many impediments to international collaboration have decreased over the last decades and, among them, the cost of travel. These changes are reflected in the overall increase in internationalization that emerges from our data (see Figure 2). During the same period, arguably, cultural distances have remained roughly constant, so that their *relative* importance, as captured by the estimated coefficient of the cross-product of distance with time, has increased.

In Europe, the overall process of integration has reduced many of the impediments to international R&D collaboration. We observe that belonging to the European Union has a strong and significant positive effect on internationalization. The results that we obtain on the impact of belonging to the Euro Zone are best seen in the light of the debate that followed the contribution of Rose (2001), who found a significant and economically very strong positive effect of currency zones on international trade. Later studies tended to confirm the presence of such an effect, but they significantly reduced its estimated size. The lively debate that followed is summarized well in Baldwin (2006), who also finds some positive effects of currency unions on trade – the Euro Zone among them. One important reason why we should be prudent in commenting on estimates of the effects of the Euro Zone is because, quite simply, it is still a relatively recent innovation and we have not observed its effects for long enough to allow for robust statistical inferences. This consideration is also the starting point for an interpretation of our results.

---

[16] A survey of senior executives representing a wide range of industries, carried out by The Economist Intelligence Unit (2007), indicated "cultural differences" as the third most important risk to developing global innovation networks (the first and the second one being, respectively, theft of intellectual property, and loss of control over innovation processes).

The estimated effects of belonging to the Euro Zone are different according to what measure of internationalization we consider. It is *negative* for the more general concept of Inventor-Applicant internationalization, and *positiv*e for our two more restrictive measures. A possible explanation of this contrast may be the following. Not all of the reasons that may lead to establishing an R&D lab abroad derive from considerations that have a strictly technological or economic nature. For example, an R&D presence abroad may also serve public relations purposes, and in some extreme cases, it may even be a precondition imposed by the local authorities in order for them to allow the establishment of a production unit. It could be that countries within the Euro Zone now find a less urgent need to set up laboratories in other countries that also belong to the currency union, and as a consequence can choose the location for their labs using criteria of choice that are strictly technological and economic in nature. More generally, we posit that beyond a certain level of functional or institutional integration, economic actors may have a less stringent need to establish their presence abroad in order to carry out a specific function.

6. Conclusions

In this paper, we have discussed the extent, and researched the causes of, internationalization of innovative activities, using a dataset on priority patent applications filed at any patent office in the European Union. We found that the degree of internationalization of innovative activities, while steadily increasing over time, is still relatively limited. Such a lasting "lack of globalization" may come as a surprise, at least if it is observed in the light of the amazing intricacy of the relations that govern today's international division of labour. In discussing our results we have stayed clear of the debate on exactly what is the nature of the emerging forms of global R&D efforts. Types of more “open” innovation, as described in Chesbrough (2003), are quite likely to be part of the story. However, adopting catchphrases to describe complex realities is risky and, in the case at hand, may conduce to confusion between the conceptually distinct ideas of “openness” and “internationalization” of innovative processes[17]. The whole issue, while certainly deserving the attention of researchers, was outside of the scope of our work.

[17] For example, for some firms, constructing an ecosystem conducive to open models of innovation may be an alternative to more traditional off-shoring of tightly controlled R&D labs. That is, in some cases we may observe more openness leading to less internationalization. Evidence

We studied the factors that determine the observed level of internationalization using a gravity model. We found that sharing a common language positively affects the degree of collaboration, as does being a member of the European Union. It is not possible to disentangle the relative role of the different ingredients that together define the European Union. Obvious candidates would be the presence of a common market, the presence of a common regulatory framework, and innovation policy instruments that encourage collaboration among actors residing in different EU countries.[18]

It also emerged that being part of the Euro Zone affects the extent of internationalization. In the case of the more restrictive inventor and applicant internationalization, the influence is estimated to be positive, although not very strong. This evidence adds to literature that, over the last few years, has tried to estimate the effect of currency unions on international trade, and it suggests that a currency union may have a positive effect on other types of international interactions besides trade. However, surprisingly, we also found evidence of a *negative* effect of the Euro Zone on the more general inventor-applicant measure of internationalization. We proposed a tentative explanation of this apparent puzzle, which boils down to a simple observation: in some cases, the desire of economic actors to be "present" in far away places, may derive from the fact that integration is *not* high enough.

Distance unambiguously affects internationalization negatively, more so in more recent years. As it emerged in studied carried out in different contexts, we also conclude that the rumors about the death of distance have been greatly exaggerated. This finding has an important policy implication in Europe. One of the explicit policy goals of the European Commission is to encourage the formation of a "European research area".[19] In order to avoid a situation where the desired type of international integration only arises between

---

on emerging models of open innovation, mostly seen from the point of view of internationalization, is in OECD (2008).

[18] An example of a European policy instrument that *forces* the collaboration of innovators from different countries of the European Union is the Seventh Framework Programme.

[19] See for example the home page of European Commission's DG Research (the Directorate General responsible for Science, Research and Technology), declaring that "European Research, and more specifically the creation of a European Research Area, are now high on the policy agenda in Europe". (http://ec.europa.eu/research/index.cfm?pg=why&lg=en . Last accessed on 1 December 2008).

countries that share a border and a language, or are relatively close to one another, innovators in member states that are relatively far apart must be provided with stronger incentives to collaborate. This does not happen currently.

We have estimated the role of distance, and we have posited that broadly defined cultural differences, which correlate with geographic distance, may play a role in explaining our findings. Obtaining a better understanding of *what* aspects of distance hinder the internationalization of R&D activities, besides being interesting in its own right, would help guide the policy maker.

References

Anderson, James (1979) “The Theoretical Foundation for the Gravity Equation” American Economic Review 69, pp. 106-116.

Archontopoulos, Eugenio; Guellec, Dominique; Stevnsborg, Niels; van Pottelsberghe de la Potterie, Bruno and van Zeebroeck, Nicolas (2007) "When small is beautiful: Measuring the evolution and consequences of the voluminosity of patent applications at the EPO", Information Economics and Policy, Vol 19, Issue 2, pp. 103-132.

Baldwin, Richard (2006), “The Euro's Trade Effect (2006), Proceedings of June 2005 workshop on what effects is EMU having on the Euro Area and its member countries” European Central Bank

Berthelon, Matias and Freund, Caroline (2008) on the conservation of distance in international trade, Journal of International Economics 75, pp. 310-320

Blum, Bernardo S. and Goldfarb, Avi (2006), Journal of International Economics, 70, pp. 384-405

Carlsson, Bo (2006) “Internationalization of innovation systems: A survey of the literature”. Research Policy 35, pp. 56-67

Ceglowki, Janet (2006) “Does Gravity Matter in a Service Economy?” mimeo

Chesbrough, Henry (2003), Open Innovation, Harvard Business Press, Cambridge, Massachusetts

Cohen, Wesley M. and Levinthal, Daniel A. (1989), “Innovation and learning: The two faces of R&D” The Economic Journal, Vol. 99, N. 397, pp. 569-596.

De Rassenfosse, Gaétan, and van Pottelsberghe de la Potterie, Bruno (2007) “Per un pugno di dollari: A first look at the price elasticity of patents” Oxford Review of Economic Policy, 23(4), pp. 588-604

Dernis, Helene; Guellec Dominique and van Pottelsberghe de la Potterie, Bruno (2001) "Using patent counts for cross-country comparisons of technology output", STI Review, 27, OECD, pp. 129-146.

Dernis, Helene and Khan, Mosahid (2004), Triadic patent families methodology, OECD, Directorate for Science, Technology and Industry, DSTI/DOC (2004)2

Disdier, Anne-Célia, and Head, Keith (2008), “The Puzzling Persistence of the Distance Effect on Bilateral Trade” The Review of Economic and Statistics, 90(1): pp. 37-48

European Patent Office (2008a), PATSTAT database, April 2008 Edition

European Patent Office (2008b), Data catalogue for the EPO Worldwide Patent Statistical Database, April 2008 Edition

Griliches, Zvi (1990) "Patent Statistics as Economic Indicators: A Survey," Journal of Economic Literature, Vol. 28, pp. 1661-1707

Guellec, Dominique, and van Pottelsberghe de la Potterie, Bruno (2001), "The internationalization of technology analyzed with patent data" Research Policy 30, pp. 1253-1266

Heckman, J. J. (1979) "Sample Selection Bias as a Specification Error" Econometrica 47, pp. 153-161.

Jaffe, Adam B. and Lerner, Josh (2004), "Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What to Do About It" Princeton University Press, Princeton, PA.

Kimura, Fukunari and Lee, Hyun-Hoon (2006) The Gravity Equation in International Trade in Services, mimeo

Le Bas, Christian and Sierra, Christophe (2002) Location versus home advantages' in R&D activities: some further results on multinationals' location strategies, Research Policy 31, pp. 589-609.

Lewer, Joshua J, and Van den Berg, Hendrik (2007), "A gravity model of immigration" Economic Letters 99, pp. 164-167.

OECD (2008) Open Innovation in Global Networks, Paris

Patel, Pari and Pavitt, Keith (1991), "Large firms in the production of the world technology: An important case of "non-globalization" Journal of International Business Studies 22 (1), pg. 1-21

Patel, Pari and M. Vega (1999) Patterns of internationalisation of corporate technology: location versus home country advantages. Research Policy 28 (2/3), pp. 145-155

Peri, Giovanni (2005) Determinants of Knowledge Flows and Their Effect on Innovation, the Review of Economics and Statistics, 87(2): pp. 308-322

Rose, Andrew (2000), One Money, One Market: Estimating the Effects of Common Currencies on Trade," Economic Policy, Vol. 15 (April), pp. 7-45

Smith, Keith (2005), Measuring Innovation, in Fagerberg, Jan; Mowery, David C. and Nelson, Richard R. (ed.), the Oxford Handbook of Innovation, Oxford University Press, Oxford, U.K.

The Economist Intelligence Unit (2007) Sharing the Idea: The Emergence of Global Innovation Networks, London.

Tables

Table 1. Fractional counts of three fictitious patents

I	II			III			IV	
	$Inv_{US,p}$	$Inv_{DE,p}$	$Inv_{FR,p}$	$App_{US,p}$	$App_{DE,p}$	$App_{FR,p}$	$\sum_{i=1}^{N} Inv_{ip}$	$\sum_{i=1}^{N} App_{ip}$
P=1: Inv: DE, FR, US, US	0.5	0.25	0.25				1	
P=1: App: DE, US				0.5	0.5	0		1
P=2: Inv: DE, DE, FR, FR	0	0.5	0.5				1	
P=2: App: FR, US				0.5	0	0.5		1
P=3: Inv: FR, US	0.5	0	0.5				1	
P=3: App: US, US				1	0	0		0
$Inv_i = \sum_{p=1}^{P} Inv_{ip}$	**1**	**0.75**	**1.25**					
$App_i = \sum_{p=1}^{P} App_{ip}$				**2**	**0.5**	**0.5**		

Table 2. Computation of measures of internationalization of three fictitious patents

$InvInv_{ij} = \sum_{p=1}^{P} Inv_{ijp}$	j = US	j = DE	j = FR	$\sum_{j=1}^{N} InvInv_{ij} = Inv_i$
i = US	0.5	0.125	0.375	**1**
i = DE	0.125	0.3125	0.3125	**0.75**
i = FR	0.375	0.3125	0.5625	**1.25**
$\sum_{i=1}^{N} InvInv_{ij} = Inv_j$	**1**	**0.75**	**1.25**	

$AppApp_{ij} = \sum_{p=1}^{P} App_{ijp}$	j = US	j = DE	j = FR	$\sum_{j=1}^{N} AppApp_{ij} = App_i$
i = US	1.5	0.25	0.25	**2**
i = DE	0.25	0.25	0	**0.5**
i = FR	0.25	0	0.25	**0.5**
$\sum_{i=1}^{N} AppApp_{ij} = App_j$	**2**	**0.5**	**0.5**	

$InvApp_{ij} = \sum_{p=1}^{P} InvApp_{ijp}$	j = US	j = DE	j = FR	$\sum_{j=1}^{N} InvApp_{ij} = Inv_i$
i = US	0.75	0.25	0	**1**
i = DE	0.375	0.125	0.25	**0.75**
i = FR	0.875	0.125	0.25	**1.25**
$\sum_{i=1}^{N} InvApp_{ij} = App_j$	**2**	**0.5**	**0.5**	

Table 3. Type of applicant vs. type of internationalization. Sample of 300 international patents

	Type of internationalization			
		InvApp and:		
Type of applicant	**Only InvApp**	**Only InvInv**	**Only AppApp**	**InvInv & AppApp**
Firm	10.29%	13.59%	22.22%	15.38%
MNE HQ	67.43%	55.34%	44.44%	53.85%
MNE SUB HQ	17.71%	12.62%	27.78%	115.38%
MNE SUB	4.57%	5.83%	5.56%	-
Person	-	1.94%	-	-
Public Res Inst	-	7.77%	-	7.69%
University	-	2.91%	-	7.69%
Sums by column	100%	100%	100%	100%
% in sample (n=300)	58.33%	34.33%	3.00%	4.33%
% in population (N=72291)	54.26%	36.30%	3.86%	5.57%
Pearson χ^2 (18 degrees of freedom) = 35.09 P-value = 0.009				

Legend:

Type of applicant: A "Firm" is any firm, which is not a multinational corporation, or does not belong to one. MNE HQ: the Headquarter (HQ) of a MNE. MNE SUB: the subsidiary of a MNE which is registered in a country different from HQ. MNE SUB HQ: the subsidiary of a MNE which is registered in the same country as the HQ. Person: a physical person. Public Res Inst: Research institute that are financed mostly by public administrations. University: Universities.

Type of internationalization: see Section 2.

Pearson χ^2 test on the independence of the two characters. Value of the test statistic and P-value are relative to the null hypothesis of independence.

Table 4. Type of applicant vs. Patent office of first filing. Sample of 300 international patents

	Patent office of first filing		
Type of applicant	**EPO: never**	**EPO: first filing**	**EPO: first filing and/or later**
Firm	13.04%	9.68%	12.26%
MNE HQ	60.87%	63.44%	62.26%
MNE SUB HQ	14.01%	21.51%	15.09%
MNE SUB	4.83%	5.38%	43.77%
Person	0.97%	-	0.94%
Public Res Inst	4.35%	-	3.77%
University	1.93%	-	1.89%
Sums by column	100%	100%	100%
% in sample (n=300)	33.66%	31.00%	35.33%
% in population (N=72291)	35.16%	31.10%	33.73%
a) Pearson χ^2 (6 degrees of freedom) = 9.66 P-value = 0.140			
b) Pearson χ^2 (6 degrees of freedom) = 4.60 P-value = 0.596			

Legend:
Type of applicant: see note in Table 1 above
Type of office of first filing: EPO: European Patent Office; Not EPO: any other EU27 national patent office. Pearson χ^2 test on the independence of: a) Ho: Type of applicant independent of "EPO never vs. EPO first filing"; b) Ho: Type of applicant independent of "EPO never vs. EPO first filing and/or later" the two characters. Value of the test statistic and P-value are relative to the null hypothesis of independence.

Table 5. Regression results: InvApp

Dep variable: **ln(** $InvApp_{ij}$ **)**	**All countries**		**EU25 only**	
	OLS	**Heckit**	**OLS**	**Heckit**
ln(agginv_x)	.0785 0.116	.2394 0.000	.0768 0.796	.2110 0.008
ln(aggapp_y)	.1672 0.007	.2851 0.000	.3057 0.001	.4135 0.000
ln(dist)	-.2594 0.000	-.4674 0.000	-.0401 0.653	-.2935 0.001
ln(dist)*time	-.0071 0.077	-.0074 0.064	-.0321 0.000	-.0279 0.000
border	.5880 0.000	.6830 0.000	.4782 0.000	.5490 0.000
language	1.0808 0.000	1.2212 0.000	1.2022 0.000	1.4158 0.000
EU Union	.0797 0.259	.0702 0.304	.5757 0.000	.5060 0.000
Euro Zone	-.1410 0.061	-.1740 0.019	-.0589 0.467	-.1160 0.150
Techproximity	.4493 0.000	.4561 0.000	.6945 0.000	.7562 0.000
n	3958	3958	2448	2448
R^2	0.6157		0.6334	
λ		.7831		.6797
Distance elasticity, 1990 and 2004:		-0.471, -0.603		-0.319, -0.816

Note for Tables 5, 6 and 7.

λ expresses the relative importance of the truncation of the dependent variable in the Heckman (1979) estimator.

Table 6. Regression results: InvInv

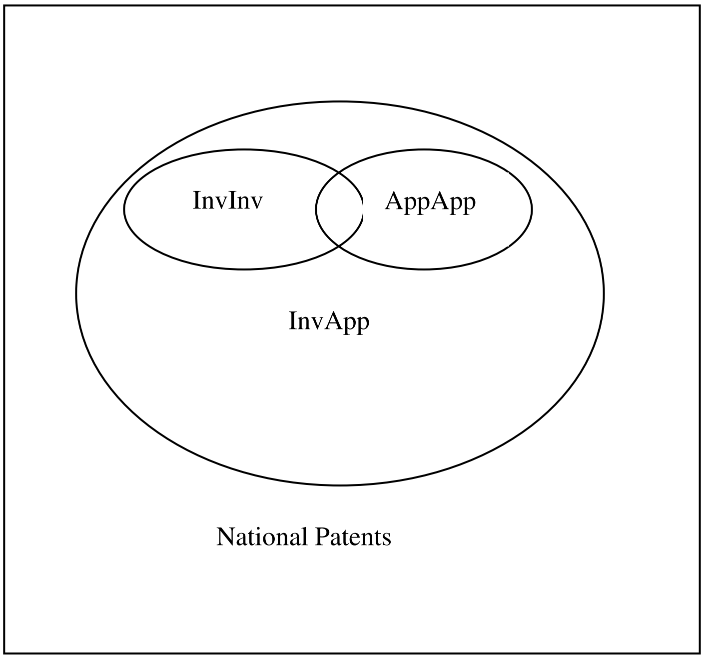
Dep variable: $\ln(Inv_{ij})$	All countries		EU25 only	
	OLS	**Heckit**	**OLS**	**Heckit**
ln(agginv_x)	.2012 0.000	.3180 0.000	.0800 0.155	.1800 0.005
ln(agginv_y)	.2012 0.000	.3180 0.000	.0800 0.155	.1800 0.005
ln(dist)	-.2300 0.000	-.3148 0.000	-.0952 0.169	-.2419 0.000
ln(dist)*time	-.0053 0.090	-.0077 0.016	-.0031 0.000	-.0265 0.000
border	.8906 0.000	.9557 0.000	.6368 0.000	.7026 0.000
language	1.0424 0.000	1.1681 0.000	1.4030 0.000	1.5748 0.000
EU Union	.2741 0.000	.2535 0.000	.4906 0.000	.4552 0.000
Euro Zone	.06181 0.305	.0151 0.800	.2193 0.001	.1548 0.015
Techproximity	.1477 0.061	.1579 0.046	.3415 0.001	.3689 0.000
n	4276	4276	2726	2726
R^2	0.6534		0.6792	
λ		.5470		.5155
Distance elasticity, 1990 and 2004:		-0.320, -0.458		-0.267, -0.739

Table 7. Regression results: AppApp

Dep variable: $\ln(App_{ij})$	All countries		EU25 only	
	OLS	**Heckit**	**OLS**	**Heckit**
ln(aggapp_x)	.1057 0.102	.2502 0.000	.0322 0.689	.1203 0.150
ln(aggapp_y)	.1057 0.102	.2502 0.000	.0322 0.689	.1203 0.150
ln(dist)	-.0123 0.822	-.2238 0.000	.3127 0.000	.0528 0.542
ln(dist)*time	.0048 0.231	.0047 0.223	-.0259 0.001	-.0254 0.001
border	.4763 0.000	.6871 0.000	.4536 0.000	.6077 0.000
language	.9898 0.000	1.164 0.000	1.0911 0.000	1.3676 0.000
EU Union	-.0129 0.868	-.0687 0.337	.1860 0.059	.1343 0.142
Euro Zone	.1083 0.150	.1571 0.029	.3497 0.000	.4306 0.000
Techproximity	.1216 0.293	.2304 0.050	.2821 0.061	.3290 0.028
n	2146	2146	1328	1328
R^2	0.543		0.5805	
λ		.8229		.7297
Distance elasticity, 1990 and 2004:		-0.218, -0.134		+.027, -0.427

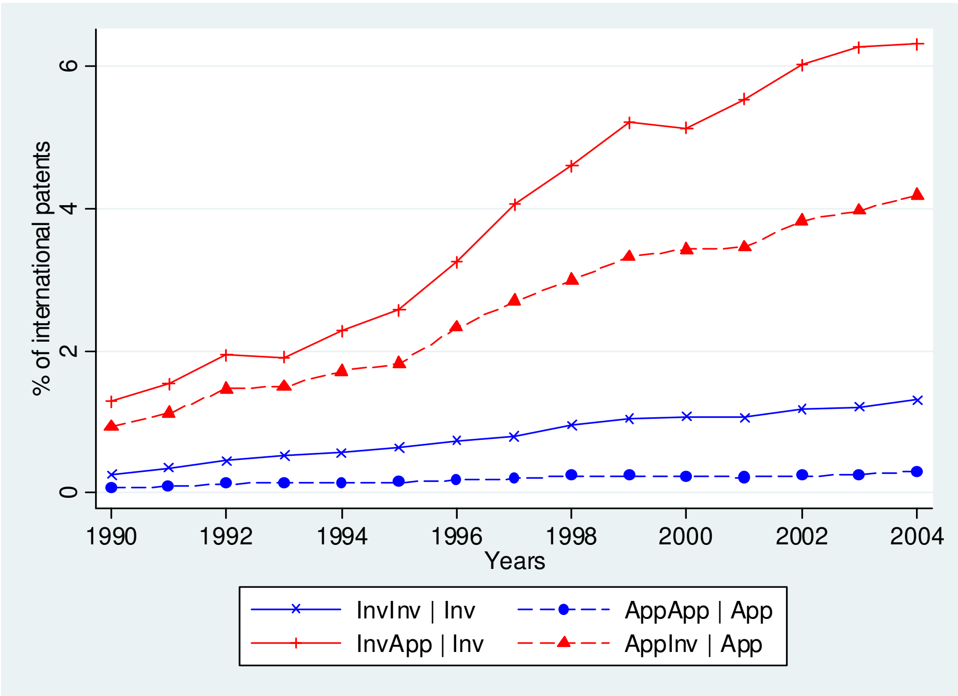
Figures

Figure 1. Relations among the different types of internationalization



Legend: Rectangle: all patents; InvApp: all international patents necessarily display (also) Inventor-Applicant internationalization

Figure 2. International patent, EU27, national averages



Note: InvInv | Inv: Eq. 8; AppApp | App: Eq. 8'; InvApp | Inv: Eq. 9; AppInv | App: Eq. 9'
Source of the data: Analysis of the Patstat database (April 2008 release) (See Section 3 for a description).