Cross Country Differences in Patent Propensity: a Firm-Level Investigation*

Francesca Lotti
Bank of Italy

Fabiano Schivardi
Bank of Italy and CEPR

February 9, 2006

Abstract

We investigate cross country differences in patent propensity in the European Union using the European Patent Office database and the database on European firms AMADEUS. We ask how much of the differences in patent propensity at the aggregate level can be accounted for by a more disaggregate analysis. First, we find that the large and persistent cross country differences cannot be explained by the sectoral composition only: patent propensity differs substantially even within sectors. Second, country effects play an important role in determining the probability that a firm engages in innovative activity, even after controlling for firms characteristics; on the contrary, they are less relevant for the amount of innovative output of those firms that do engage in patenting. This indicates that some countries might be more conducive for firms to undertake innovation activity, but that, once this is done, innovative firms are less sensitive to the home country effect.

Keywords: Patents, Manufacturing, European Patent Office, AMADEUS.
JEL Classification: L60, O32, O34.

*We would like to thank Gianpaolo Lopez and Emmanuele Somma for their help with the EP database, including the SOUNDEX algorithm for the matching, and Marco Chiurato and Elena Genito for research assistance. A previous version of the paper was presented at the conference "Nuovi Temi per la Politica Economica" jointly organized by the "Ente Luigi Einaudi" and the "Giornale degli Economisti" held in Rome in November 2005. We wish to thank the organizers, Michele Polo and Daniele Terlizzese, the audience and the discussants, Stefano Breschi and Stefano Usai, for their useful comments. Nonetheless, we are responsible for any mistake. The views expressed by the authors do not necessarily reflect those of the Bank of Italy. Emails: francesca.lotti@bancaditalia.it, fabiano.schivardi@bancaditalia.it.

1 Introduction

There is a widespread consensus that, among industrialized countries, innovation is the main engine of growth. In the policy debate, the need to foster the innovation capacity is always cited as the main channel to increase growth prospects. This is for example a central pillar of the Lisbon strategy to improve the performance of the European economies, which have lagged behind the US especially in terms of innovation activity. Even within Europe, there are large cross country differences in innovation performance. This paper aims at assessing such differences and at investigating their causes.

We use patent data from the European Patent Office (EPO) for the 15 countries of the European Union (EU15). After the seminal contribution of Schmookler (1972), who compiled many time series of patents by industry for a time span of a century, it has become clear that patent statistics, even though subject to issues of interpretation, provide a unique source of systematic information about the innovative process. Correspondingly, we use patent counts to measure innovative output. Compared to other measures of innovation, such as those obtained form firm surveys, patents have the advantage of being a clearly and unambiguously defined indicator of innovative output. However, at least two caveats should be kept in mind. First, patents are only one aspect of innovation activity. For instance, many process or incremental innovations are not worth to be patented; also, firms can use alternative modes of intellectual property protection, such as secrecy and continuous innovation.[1] Accordingly, not all the innovative activity flows in a patent. Nonetheless, previous research suggests that patents are strictly correlated with the total innovative output. For example, using US data, Pakes and Griliches (1984) find a strong relationship between R&D and the number of patents across firms and industry. Crepon, Duguet and Mairesse (1998) show that, for French firms, R&D expenditure, the number of patents and the share of sales coming from innovative products are positively correlated. A second issue is that not all patents have the same value or innovative content. However, as a first approximation, it is reasonable to use simple patent counts as a measure of innovative output.[2]

Our main objective is to determine how much of the cross country differences in patenting activity are due to some "aggregate" country attribute, such as the education of the labor force, material and immaterial infrastructures etc.[3], and how much can be explained by a more disaggregated analysis. To answer this question, we proceed in three steps. First, we study

[1] Griliches (1990) states: "Not all inventions are patentable, not all inventions are patented, and the inventions that are patented differ greatly in terms of quality".

[2] Hall, Jaffe and Trajtenberg (2001) propose the use of patent citations as a measure of the value of a patent, a possibility that we plan to exploit in future work.

[3] See Furman, Porter and Stern (2002) for an analysis of the determinants of the national innovative capacity at the aggregate level.

patent activity at the country level, analyzing the evolution of patent propensity (PP) indicators, such as patents over population. Second, we consider to what extent such divergencies are due to differences in sectoral specialization patterns. Third, we analyze PP at the level of the firm. In fact, firm characteristics can impact the overall PP at the country level: if small firms are less likely to patent, then a size structure tilted toward the small size could reduce the aggregate PP. By using firm data, we will be able to assess if cross country differences in propensities survive after controlling for individual firms' attributes.

We find that countries do differ substantially in their PP: in 2000 the number of patents per 100,000 inhabitants ranged from 0.26 for Portugal to 27 for Finland. These differences are persistent, even if some evidence of catching up emerges over the period 1980-2000. Also, we find that they are attributable to differences in both intensity (the amount of investment in innovation activity) and efficiency (measured by the number of patents per researcher).

Sectoral composition plays a minor role in determining cross countries differences in PP: most of the differences survive when keeping sectoral composition constant. This implies that the problem of the laggard countries can only partially be explained by the fact that they are specialized in productions with a lower PP. Rather, they tend to have a lower propensity in all sectors.

To study PP at the firm level, we merge the information on patents from the EPO to that on individual firms from AMADEUS, a large database of around 200,000 medium and large European companies with information on financial records and employment. We show that, while not representative of the whole economy, firms in the AMADEUS database offer a satisfactory picture of patenting behavior. Given that only a small fraction of firms are engaged in patenting, we use an Heckman selection model that distinguishes between the decision to actually engage in patenting activity and the number of patents, conditional on being engaged. Our goal is to determine if the country effects, measured by the country dummies in the regressions, explain less of the cross country variability when accounting for firm characteristics. We find that this is actually the case, but with a difference between the two stages of the regression. In the first stage, i.e. the decision to undertake innovation activity, the dummies remain significant even after controlling for firm characteristics; the effects instead basically vanishes in the second stage, i.e. in the regression of patent intensity for firms that have at least one patent. This evidence indicates that country effects are most relevant to determine the likelihood that a firm invests in innovation; conditional on doing so, the country of residence plays less of a role for the amount of innovative output produced. One interpretation is that firms that pass the patenting threshold are more internationally oriented and less sensitive to home country effects. We also find that firm size has a positive impact on the likelihood to patent, while PP, defined in terms of patents per employees, seems to be decreasing with firm size (the elasticity is far below one).

In terms of policy, our results imply that moving towards a size structure with larger firms would increase the number of firms engaged in innovation activity but not necessarily the average PP. Therefore, policies that would help firms to pass the patenting threshold, for example through tax incentives, grants or technical assistance for newly patenting firms, could have a large impact on country level patent activity, generating an incentive to undertake more R&D activities.

The rest of the paper is organized as follows. Section 2 contains the country level analysis, while Section 3 the sectoral one. In section 4 we introduce the firm database and undertake the econometric investigation of PP at the level of the firm; we conclude the section with a brief discussion of the results and of their main implications in terms of innovation policy.

2 Country Analysis

In this section we examine the innovative performances of the EU-15 countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the UK) in terms of patent applications at the EPO. The main objective of this section is to give a picture of the patenting performance of EU countries at the level of the aggregate economy, regardless of sectoral composition and market structure. We will investigate differences in PP across countries, defined here as patents over population, and their evolution over time; we will then move on to analyze the role of differences in the intensity of the innovation effort at the country level (how much countries differ in R&D expenditures) and of its efficiency (how much they differ in terms of number of patents per unit of R&D input).

2.1 Cross country differences in PP

Our data on patents come directly from the filings at the EPO. The EPO was established with the aim of increasing cooperation among European countries in terms of intellectual property rights, by means of a centralized patenting procedure and harmonized rules. A patent granted by the EPO is a bundle of property rights potentially effective in every country that agreed to sign the European Patent Convention, established in 1973. This has enormously simplified the application procedures, guaranteeing homogeneity in the patent examination process and lower costs for an applicant who is seeking protection in more than one country.

When an inventor (an individual, a corporation or a public institution) wants to protect her innovation with a European patent, the procedure requires to fill an application; after that, the examination procedure starts. After 18 months the application is published, even if the patent has not been granted yet. The duration of the examination phase is not homogeneous: on average, it is much longer at the EPO than at the U.S. patent office. We use data provided directly by the EPO about patent applications, no matter if they were granted or not. We are mainly interested in measuring innovative activity of countries, industries and firms, and for this reason we take a patent application as a signal of the presence of some innovative output. We use the so called "priority date" as the date to identify patents in time: this is the date in which the first patent application occurred, and therefore closer to the date in which the innovation took place. Often a patent has more than one proprietor: in those cases, as the order of the names matters, we attribute the whole patent to the first name/country that appears in the application. Unfortunately the EPO does not provide any information about the nature of the patenting subject; it is therefore not possible to distinguish individuals from corporations or public institutions, to ascertain if countries differ in the mix of patenting subjects.

Since its foundation, in 1978, the number of patent applications at the EPO grew at a yearly average rate of 9% until 1995 and then at 14% afterwards, from 24,045 applications in 1980 to 107,758 in 2000. This sustained rate of growth is not confined to Europe: indeed, since the mid-eighties, a substantial increase in the number of applications has been recorded at the U.S. Patent Office (Hall, 2004).[4] However, the growth is stronger in Europe, suggesting that, in addition to an international trend, the applications at the EPO reflect the growing importance of European patents as a tool for intellectual property protection in a progressively more integrated European economy.

The share of patents assigned to EU members has declined over time, as Japan and the U.S. started to seek protection for their innovations into the European market (Figure 1). The same phenomenon emerged in the U.S. in the Eighties, again reflecting a progressive integration of the markets for intellectual property rights. Table 1 reports the share of applications of each EU-15 country on the total number of applications of these countries from 1978 to 2000. Germany has been by far the country with the highest share: alone, it accounts for between 40 and 49% of the total applications; France is the second major contributor, with a share ranging from 15 to 20%, followed by the UK and Italy. At the other extreme, Greece, Ireland, Luxembourg, Spain and Portugal each accounted on average for less than 1% of the total. Of course, these numbers strongly depend on the size of the economy and therefore should not be interpreted in terms of PP.

[4]The reasons of the "patent explosion" is the topic of an ongoing debate, started by the seminal contribution of Kortum and Lerner (1998).

In terms of time patterns, there is a clear tendency for countries with low initial shares to increase their contribution, suggesting that a process of convergence might be at work; at the same time, countries with a high initial value tend to maintain or only slightly decrease their share; a clear exception is the UK, whose contribution went from well above 15% to around 10% in the latest years.

Figure 2 reports the evolution of the PP, defined as the ratio of the number of patents to the population (in 100,000). For clarity, we split between low (Belgium, Greece, Ireland, Italy, Portugal, Spain, UK) and high (Austria, Denmark, Finland, Germany, France, Luxembourg, Netherlands, Sweden) propensity countries, defined as above or below the median in 2000. The figure shows that PP differs greatly across countries: in 2000 the mean is 11.5 patents for 100,000 inhabitants and the distribution ranges from a minimum of 0.26 for Portugal and 0.37 for Greece to a maximum of 25 for Sweden and 27 for Finland. This indicates that the differences of Table 1 are not only due to country size.

Within the high propensity group, there is no evidence of convergence: Austria, Denmark and France start at the bottom of the range and, by 2000, record a more pronounced gap with respect to the leading countries (Germany, Luxembourg and Sweden). The most spectacular growth is recorded by Finland, that starts out at the bottom of the distribution but, thanks mostly to the high growth rate recorded in the 90s, ends up at the top.

A similar pattern emerges for the low propensity country. The laggards (Greece, Portugal and Spain) are characterized by markedly lower values both at the beginning and at the end of the period with respect to the UK, Belgium, Italy and Ireland. Even in this group there is a fast growing country, Ireland, which records a remarkable growth since the mid nineties. Belgium is characterized by a steady growth while Italy has a more uneven profile, flat from the late 80s to the mid nineties and increasing afterwards. The UK shows a flat profile from the late 80s to the late 90s, which explains the marked reduction in its share of patents seen above.

To better assess if there is some evidence of catching up, Figure 3 plots the average annual growth rate in PP against the initial value.[5] The figure shows that there is indeed a catching up process at work: countries that started out with very low levels (Greece, Ireland, Portugal and Spain) do show higher growth rates (in the rage of 15-20%); however, their initial values are so low that they contribute to a limited amount to convergence: for example, Portugal goes from 0.01 to 0.22 patents per 100,000 inhabitants, growing at a remarkable annual rate of 23%; still, it records a very high distance from the EU15 average in 2000 (10.9). For the other countries,

[5]We have computed the initial and the final values using an average over three years (1981-84 and 1998-2000) to smooth out high frequency variations that, particularly for small countries, could add some noise to the calculations.

the evidence is less marked: Italy, Belgium and Denmark, that have lower initial values, record growth rates only marginally above the mean.

This evidence suggests that there are consistent differences in PP among countries and that such differences are persistent. While some tendency to converge emerges, both the very low initial values of the laggard countries and the modest differences in growth rates for those in the intermediate levels imply that dispersion is still very high in 2000: in fact, there is a difference of a factor of 100 between the lowest and the highest propensity country and the coefficient of variation between 1980 and 2000 has only decreased from 1 to 0.8.

2.2 Efficiency and intensity of innovation activity

The analysis so far has been based on the patents to population ratio. This statistic does not take into account the fact that countries might differ in the rate of participation to the labor market: if only employed people contribute to the innovation process, then the use of population might give a questionable picture of the PP at the country level. In reality, this is not likely to be a major issue, given that the ratio of employment to population varies much less than PP. To show that this is the case, in Figure 4 we report the ratio of patents to total employment for year 2000. The ranking is basically the same as with population, as are the differences across countries. This indicates that PP can be analyzed with either indicators; given that in the next sections we will be using sectoral and firm level data, for which PP can only be defined in terms of employment, in what follows we use the number of patents over employment as the indicator of PP.

One major determinant of PP is the amount of investment in R&D: European countries are characterized by very different investment rate in R&D. Figure 5 plots PP and the share of GDP devoted to R&D. Variations in R&D investments are substantial, ranging from slightly more than 0.5% of GDP for Greece and Portugal to around 3.5% for Finland and Sweden. The relation is strongly positive: countries that invest more in R&D do patent more. Therefore, R&D investment explains a substantial proportion of the cross country variations in PP.

While R&D intensity is clearly an important determinant of PP, the variability of the latter is more pronounced than that of the former. A second source of variation in propensity across countries is the "efficiency" of the innovation process, ie. the rate of patenting for unit of investment. We investigate the role of each of these components using the share of workers involved in R&D over total employment as a measure of intensity and the number of patents for researchers for efficiency. Figure 6 reports these variables for the year 2000. If countries differ mostly in their innovation intensity, then cross country gaps when measured in terms of patent per researcher should be substantially narrower than for overall PP. Clearly, the graph shows that there are large differences in PP even when this is computed in terms of R&D personnel. This indicates that another important determinant of cross country variation is the efficiency of the research process. Countries differ markedly also in the share of R&D personnel, but to a lesser extent: the ratio between the highest and the lowest value is around 30 for efficiency, against 5 for intensity.

To get a more precise indication of the role of efficiency and intensity, define $p_j = P_j/L_j$, $p_j^{RES} = P_j/R_j$ and $r_j = R_j/L_j$, where P_j is the number of patents of country j, L employment and R the number of researchers. Let $\bar{p}$ and $\bar{r}$ be the corresponding values at the EU-15 level. Then, the deviation of PP from the EU-15 average can be decomposed as:

$$\begin{aligned} \frac{p_j - \bar{p}}{\bar{p}} &= \frac{p_j^{RES} r_j - \bar{p}^{RES}\bar{r}}{\bar{p}} = \\ &= \frac{(p_j^{RES} - \bar{p}^{RES})\bar{r}}{\bar{p}} + \frac{(r_j - \bar{r})\bar{p}^{RES}}{\bar{p}} + \frac{(p_j^{RES} - \bar{p}^{RES})(r_j - \bar{r})}{\bar{p}} \end{aligned} \tag{1}$$

The first term in the second line represents the deviation of PP from the EU-15 mean due to the difference in efficiency, the second in intensity and the third is a cross term which is positive when deviations of efficiency and intensity have the same sign. Results from this decomposition are reported in Table 2. In general, both effects are important for explaining cross country differences. Efficiency plays a more prominent role for Belgium, Spain, Greece and Portugal (mostly the laggard countries), while intensity is more important for Finland, whose prominence is only due to this factor, and Sweden. For all other countries both effects are relevant: for example, Italy's lower propensity is attributable in fairly equal part to efficiency (-28%) and intensity (-36%), as is the case for Germany (51% due to efficiency and 29% to intensity). The cross terms is positive in 10 cases out of 15, indicating that countries with a higher efficiency also tend to have a higher intensity and vice versa. The noticeable exceptions are Belgium, Denmark and France, where an efficiency far below the mean is accompanied with a substantially stronger intensity: these countries seem to compensate a lower productivity of the innovation activity with a higher effort.

We have controlled for the robustness of our results using R&D expenditure (Figure 7). According to this indicator, Germany and the Netherlands have the lower cost of patent production, while Spain, Greece and Portugal the highest. The ranking is basically unchanged with respect to the number of researchers. We are therefore confident that the conclusions above are robust with respect to the different measures of innovative inputs.

3 Sectoral analysis

Having established that countries differ systematically and persistently in PP, we now move one step further in investigating the causes. One possibility is that differences are attributable to specialization patterns. It is well known that EU countries differ substantially in the sectoral composition of the economy, with Mediterranean countries more specialized in traditional, low tech sectors than northern European ones. Given that different sectors are characterized by different levels of innovation opportunities (Pavitt 1984), and therefore of PP, sectoral specialization could then be an important source of cross country differences.

To analyze the role of sectoral specialization, we restrict our attention to manufacturing, because we were able to attribute patents only to manufacturing sectors. Upon application for protection, patents are assigned a code which helps examiners and lawyers in grant and litigation processes. According to the International Patent Classification (IPC), patents are classified in terms of the technological field and not by the sector in which the firm-inventor operates. We use the IPC-ISIC concordance table developed by Verspagen et al. (1994) to link each patent to the industrial sector in which innovation took place, which clearly maps patent product or process categories into the economic sectors responsible for their creation.[6]

Table 3 reports the number of patents per 1,000 employees for each manufacturing sector for the EU-15 countries, excluding Greece and Luxembourg due to data availability. In the last column we report the total for the area; all other entries are expressed as ratio to this total, so that a value above one indicates that the PP of a country in that sector is larger than the aggregate and vice versa.

In terms of sectoral propensity, the column for the total of the area shows that low tech sectors have a lower PP, with less than 0.5 patents per 10,000 employees. The highest propensity is in office machinery and computers (10 patents per 1,000 employees), followed by pharmaceutical (7.6) and precision instruments (6.9). Between the lowest (textile and leather) and the highest (computers) there is a difference of a factor of 60, which indicates that sectoral specialization could actually play an important role in explaining cross country differences in PP.

If sectoral specialization were the main driver of cross country differences in PP but, within sectors, countries had fairly similar propensities, then the entries in the table should be around 1. This is clearly not the case: on the contrary, the table shows that there are clear country

[6]Other concordance tables are available, like the OECD Technology Concordance table (OTC), developed by Johnson (2002). While the OTC table allows a more disaggregate industrial classification, it generates, for each patent, a series of couples of sectors where, with a certain probability, the patent was generated and used. Since we are only interested in the production of patents, we stick to the simpler IPC-ISIC concordance table developed by Verspagen et al. (1994).

patterns: countries with a low overall propensity also tend to replicate the same pattern within each sector and vice versa for high propensity countries. In particular Finland, Germany, the Netherlands and Sweden are characterized by values almost always larger than 1. On the other side, Italy has a sectoral PP uniformly close to 0.5, with no value larger than 1; interestingly, the distance from the EU aggregate is lower for the traditional sectors (textile and leather, wood) and for motor vehicles; it is highest for pharmaceutical and communication equipment. Spain and particularly Portugal have very low values across the board, again with no sector with a PP that matches the area averages.

The cross-sectoral dispersion within each country is another indicator of the relevance of an overall country effect. If sectoral propensities are very dispersed within countries, then we would find support for a Marshallian view of the innovation process: countries specialize in some productions in which they are relatively more efficient compared to other productions. The dispersion varies significantly across countries: the outlier is Ireland, due to an abnormal value in the motor vehicles sector; Finland also shows a high degree of dispersion, driven mostly by the high relative propensity in the various machinery sub-sectors. This country seems to fit the specialization story. Dispersion is lower for the other countries; the five biggest economies all have values between 0.3 and 0.4. The relatively low dispersion lends further support to a general country effect on propensity, independent from specialization patterns. In fact, differences in PP remain strong even at a fairly disaggregated sectoral level.

Some further insights can be gained by considering dispersion of sectoral PP across countries. The coefficient of variation varies significantly across sectors. A clear pattern does not emerge, but overall the correlation between dispersion and average PP is positive (0.35), indicating that dispersion is higher in sectors with a high PP. This suggests that cross country differences are more marked in high tech sectors.

The analysis so far indicates that country patterns might be a more prominent factor in explaining differences in PP than sectoral specialization. To properly assess this assertion, we undertake a shift and share decomposition, similar to the one of the previous section. More precisely, define as before p_j as the total number of patents over total employment for country j, p_{ij} as the number of patents over employment in sector i and country j, ω_{ij} as the share of sector i employment over total employment and the barred variables as the corresponding values at the EU-15 level. Then we decompose the deviation in country-levels PP from the EU-15 aggregate in a component due to the within sector propensity (first term in equation 2 below), one due to sectoral specialization (second term) and the interaction term (third term) as follows:

$$\frac{p_j - \bar{p}}{\bar{p}} = \frac{\sum_i \omega_{ij} p_{ij} - \sum_i \bar{\omega}_i \bar{p}_i}{\bar{p}} =$$
$$= \frac{\sum_i (p_{ij} - \bar{p}_i)\bar{\omega}_i}{\bar{p}} + \frac{\sum_i (\omega_{ij} - \bar{\omega}_i)\bar{p}_i}{\bar{p}} + \frac{\sum_i (p_{ij} - \bar{p}_i)(\omega_{ij} - \bar{\omega}_i)}{\bar{p}} \tag{2}$$

Results from this decomposition are reported in Table 4. For most countries, the main determinant of deviation from the average is the within sector propensity. This explains most of the deviation for Finland, Germany, Italy, the Netherlands, Portugal, Spain, Sweden and the UK, indicating that differences in PP are explained only marginally by the sectoral composition. The main exception is Ireland, which shows a large contribution of the sectoral composition: indeed, employment data show a strong specialization of this country in the electrical and communication machinery and computer industry. Ireland has also a large negative contribution from the interaction term, indicating that, even if specialized in high propensity sectors, the PP is lower than the average. Italy's PP is almost half of the EU-15 aggregate. Of this, only 10% is explained by the sectoral composition effect. From this respect, Italy is different from Portugal and Spain, for which, in addition to a lower propensity within sector, also sectoral composition plays an important role.

This analysis shows that sectoral composition plays a less important role in determining PP across countries. The more relevant differences come from the different propensity within sectors, indicating that countries tend to differ in systematic ways even when comparing the same sectors. In the next section we investigate how much of these differences can be accounted for by differences in industrial structures.

4 Firm Level Analysis

4.1 Data

In the light of the results depicted in the previous section, in this paragraph we bring the analysis to a further degree of disaggregation, considering PP at the firm level. We will try to ascertain how much of the differences in PP across countries are attributable to some general characteristics of each country and how much to the fact that the industrial structures are different across them. For example, it is known that firm size structures differ substantially across European countries even within sectors (Pagano and Schivardi 2003). If PP differs according to firm size, then different size structures might be responsible for the different PPs discussed in the previous sections. Stated differently, we will try do determine if such differences persist even after controlling for firms attributes.[7]

The firm level data come from AMADEUS (Analyze Major Database from European Sources), a database collected by Bureau van Dijk. AMADEUS contains balance sheet data for 200,000 non financial companies from 36 European countries. Balance sheet data are standardized according to the best accounting practices. To be included in this database, firms must satisfy at least one of the following criteria: (i) operating revenues of at least 10 million euros; (ii) total assets of at least 20 million euros; (iii) more than 100 employees. These are necessary but not sufficient criteria for inclusion: in fact, the representativeness of the sample varies from country to country. As a general rule, these criteria imply that the sample is tilted toward medium and large size firms, which makes the average company in AMADEUS more likely to be engaged in patenting activity than the average one in the population.

As before, we restrict our analysis to the EU-15 countries. The release of AMADEUS we use contains data from 1999 to 2004 for 115,686 companies in these countries. We only keep firms with unconsolidated balance sheet data, for two reasons: first, many multinational corporations have the headquarters in one country but may file the patents through their subsidiaries in other countries (mainly located in Germany and in the Netherlands); second, we want to avoid imputing the consolidated employees or sales from the affiliated firms in different countries to the one where the headquarters are located.

Assigning to each firm its patents requires extreme care, as the EPO database contains only the names of the applicants. These can be individual, public organizations or corporations, but the EPO does not provide a direct way to categorize them. For firms, the only information available is the name, address and country: we choose as firm identifier the couple (*name*, *country*), as this information is present also in AMADEUS. Unfortunately, the names reported in the EPO database are not always the exact spelling of the official names. To maximize the number of matches, we have proceeded in 3 steps, each performed recursively on the firms of each country. First, we simply matched the entire company name, after having removed all non letter characters; second, among those not combined yet, we have matched the first 15 characters of the company names; third, for the remaining firms, we used SOUNDEX, an algorithm that produces matches for strings using a weighting scheme, according to which each component of the string is assigned a certain weight and matches are produced accordingly. The last two steps resulted in some incorrect matches; therefore, we have inspected all these matches one by one. Eventually,

[7]A recent literature has stressed the role of local externalizes for innovative performance. According to this literature, geographical clusters of firms would be the most appropriate unit of analysis for the innovation process than single firms (Audretsch and Feldman 2004). For example, Paci and Pigliaru (2002) show that regional innovation propensity depends on that of the surrounding regions. Spatial clustering could then be a further element explaining cross country difference in PP. We confine these issues to future work.

we found 4,654 AMADEUS companies with at least one patent in the EPO database. Given that we have controlled each match, this procedure is unlikely to produce wrong attributions; on the contrary, most likely some firms which did apply for a patent might have not been identified, so that we should expect some degree of under-representation of patenting companies. Unfortunately, given that the EPO database does not report the legal status of the applicant, it is not possible to count how many applications are attributable to firms and therefore to come up with an estimate of the degree of under-representation.

Due to lags in updating, the EPO data on patents are reliable up to 2000. We therefore select this year for the analysis that follows. We have a total of 74,438 firms, 1,856 of which have at least one patent application in that year. A serious concern is the representativeness of the AMADEUS database, both in general and especially in terms of PP. The database might give distorted PP for two reasons. On one side, the database is tilted toward large firms, that have a higher PP (Bound et al. 1984; Cohen et al., 1987); moreover, the firm coverage is different across countries, thus making cross country comparisons problematic. On the other side, the fact that the matching procedure is not likely to identify all applying firms implies a downward bias in PP, which is likely to be uniform across countries.

Table 5 reports a comparison of the AMADEUS data with the population ones in terms of firms and employment. The coverage of AMADEUS is low in terms of firms: around 1% or less of the total number of firms are present. It is much higher in terms of employment, often above 30%, given that the firms in the database are larger than the average in the population. In fact, the average size of AMADEUS firms ranges from 151 for Belgium to 1,123 for Germany, against a population value for the whole area of 6.3. This difference is due to the fact that AMADEUS excludes basically all very small firms, that constitute the bulk of number of firms but not of employment (Bartelsman, Scarpetta and Schivardi 2005). Employment coverage is lower for Austria, Germany, the Netherlands and Portugal, all below 10%. Notwithstanding these shortcomings, the database is fairly representative of the sectoral composition of the economy: the share of manufacturing employment is in fact close to the corresponding population one for all countries. Most importantly, the PP in AMADEUS is very similar to the population one (last two columns): the total for the area is the same (0,52 patents per 1,000 employees); the values are also very similar for most countries and the country ranking is almost identical. The main differences are Ireland, for which the patent propensity in AMADEUS is substantially lower than the population one, arguably because the National Account data are incomplete in terms of employment (see the note to the table); Finland and especially Germany, for which the population propensity is lower than the AMADEUS one. For the latter, the reason is that the AMADEUS size cutoff is particularly high for this country: the average German firm in AMADEUS is five time as large as the one for the EU-15 aggregate, while the ratio is 2 for the population average. Despite these differences, the AMADEUS data give an overall picture that is close enough to the actual one to make the analysis of cross country differences in PP possible.

We have done some further data cleaning. In sectoral terms, we have excluded agriculture and public services (ateco 1-10 and ateco 76-99); we have also dropped those sectors that, in 2000, had less than 10 patents applications in total, ie. retail trade (ateco 52), hotels and restaurants (55), transport and storage (61-63), insurance (66), real estate and renting (70-71). We have also excluded Philips and Unilever and all their subsidiaries, as we were not able to attribute the large number of patents of these companies to the various subsidiaries. After this procedure, we were left with 52,652 observations. Table 6 reports descriptive statistics of the regression variables for firm observations included in the regressions, separately for each country and distinguishing between firm with and without one patent application in 2000. For the EU-15 aggregate, patenting firms are substantially larger than non patenting ones both in terms of employment and total assets; the same holds for almost all countries. No clear patterns instead emerge for operating revenues per employees or for the ratio between intangible and fixed assets. The average number of patents for patenting firms is 8, with a maximum of 31 for Germany. PP conditional on patenting is one order of magnitude larger than when considering all firms. This clearly indicates that PP indicators are based on the contribution of a relatively small number of firms that, on average, patent substantially above the EU15 average. In terms of number of observations, Greece, Ireland, Luxembourg and Portugal have 3 or less patenting firms, so that the results for these countries are statistically weak.

4.2 Testing Strategy

Our econometric model evaluates how firm characteristics can affect the likelihood of patenting, both in terms of firm PP and of patent count. In particular, we want to ascertain if the role of country dummies changes after controlling for firm characteristics: we will ask if laggard countries are characterized by the prevalence of firms with features not conducive to patenting (for example size) or by firms that patent less in the same conditions.

The dependent variable exhibits a massive presence of zeros, as we observe the number of patents only for patenting firms. In order to account for this issue, we use a sample selection framework, in which a firm first decides if to apply for any patent at all, and then, conditional on having decided to apply, how many patents to apply for. We use this modelling strategy because there is widespread evidence that many firms do not engage in R&D activity (see for example the *Community Innovation Survey*, run by Eurostat). In fact, it seems plausible that the decision to engage in innovative activity is governed by a different process from that determining how much activity to undertake, conditional on being engaged. For example, managers without scientific training might be less prone to set up a research unit but, once this is done, they might become accustomed with the innovation process and behave accordingly.

The natural framework to model this process is Heckman (1979) selection model. Ideally, we would need a further variable determining if the firm is engaged in patenting activity, such as a specific question in a survey or the R&D expenditure. Unfortunately, we do not have this information in our data. We will therefore use the zero patent application outcome as an indicator that the firm is not engaged in patenting. One should keep in mind that this might also signal that a firm has been engaged in the activity, for example because the research project failed. We have no way to determine how relevant this possibility is; it should be noted that, by considering applications and not grants, we are not misclassifying cases in which the firm does apply but the patent is denied by the EPO. In any case, results should be interpreted keeping this caveat in mind.

Let p be a random variable denoting our variable of interest (the PP or the number of patents) and z an indicator variable that takes value 1 if the firm applies for at least one patent. The statistical model can be expressed as:

$$\begin{cases} p_i = \mathbf{x_i}\beta + u_i \\ z_i = 1\left[\mathbf{w_i}\delta + v_i > 0\right] \end{cases} \tag{3}$$

where $\mathbf{x}$ and $\mathbf{w}$ are observable firms' characteristics and p is observed if and only if z=1. The first stage is probit model

$$Pr\left(z = 1\right) = G\left(\mathbf{w}\delta\right) \tag{4}$$

with $G\left(z\right) = \int_{-\infty}^{z} \phi\left(v\right) dv$ and $\phi\left(\cdot\right)$ is the standard normal density. The second stage is an OLS regression in which p is, alternatively, the patents/employees ratio (PP) or the log of patent count, augmented with the Mills ratio obtained from the first stage regression.

We will include as explanatory variables country dummies, 28 two digit level sector dummies and the firm information obtainable from AMADEUS. We assume that PP depends on size, firms' efficiency (measured by operating revenues over turnover) and asset composition (the ratio between intangible and fixed assets). As an exclusion restriction, we use a dummy which is equal to 1 if the firm ever applied for a patent before year 2000. This modelling strategy can be rationalized by assuming that there exists a fixed cost when deciding to file the first application.

For example, a firm might have to learn the current regulation and filing procedures. We therefore assume that, if a firm has already applied for a patent in the past, it has a lower cost of learning the application procedures this year; at the same time, conditional on filing at least one application in the current year, it has learned such procedures, so that the past application dummy should not influence the current number of applications.

Table 7 reports the results of the first stage regression of the Heckman procedure. The excluded country is Germany: in fact, this is the country with the highest patent per firm (Table 5), so that the country dummies can readily be interpreted as distance from the "leader". We report the marginal effects of the coefficients, so that each country dummy coefficient indicates the percentage difference from Germany in the probability that a firm applies.

The first column reports the estimates of a model with only the country and sector dummies. As expected, all dummies have a negative sign and are significant at conventional levels, with the exception of Austria and Luxembourg, confirming that the German firms included in our sample have the highest probability to patent. This could of course be a consequence of the different sampling rules applied by AMADEUS. Still, the effects are sizeable, between 0.2 and 0.5%, compared with a predicted probability of patenting at the sample mean of the observations of 0.4 and an actual one of 2.4%. The distance is higher for Italy and Spain, in line with the previous analysis; it is also substantial for France, the Netherlands and the UK (0.4%). The dummy for previous applications has a very large value (16%) and is highly significant.

In the next two columns we include firm specific controls. Model 2 includes the log of employment, the log of operating revenue over total assets and the ratio of intangible to fixed assets. All variables have the expected sign: larger, more profitable firms and those with a larger share of intangible assets are more likely to apply. The country dummies remain mostly significant, but with a smaller magnitudes in absolute value. A similar conclusion is reached in Model 3, that adds a quadratic in size. According to this specification, the probability of having at least one patent application is non monotone in size: it first declines and then increases. However, the decreasing portion is rather narrow: the probability of applying for a patent reaches a minimum at 14 employees, which is around the 10^{th} percentile of the size distribution. Similar results are obtained in the last model, where size is measured by total assets: also in this case we obtain a non monotone effect of size when including the quadratic.

As stated above, the country comparisons are likely to be influenced by differences in selection rules across countries, even if Table 5 was reassuring in this respect. However, even if there were large effects on the *levels* of the estimates, it should not be so on their *changes* when we include firm characteristics. Indeed, it turns out that, even with firms' controls, most of the country dummies remain statistically significant and sizeable in magnitude. To give a better appreciation of the country effects, in the top panel of figure 8 we report the estimates of the coefficients on the country dummies for the specification of Model 1, with no firm controls, and of Model 3, with the quadratic in employment. The graph clearly shows that, when controlling for firm characteristics, the differences with German firms in the probability of patenting decrease for all countries; at the same time, the drop is limited: the sum of the coefficients on country dummies goes from -5.1% to -3.7%. This indicates that, even after controlling for firm characteristics, the country effect still plays an important role in explaining the patenting probability.

We now turn to the second stage of the analysis, i.e. the PP conditional on having at least one application. We first consider the log of the number of applications as the dependent variable and use the same regressors as in the previous table, with the exclusion of the previous patents dummy. Results are reported in Table 8. Again, we begin with a specification with only the country and sector dummies. The coefficient of country dummies are negative and significant for all countries with the exception of the Netherlands and the 4 countries with few observations mentioned above. The values are on average -0.5, indicating that German firms tend to patent 50% more than those of the other countries. Compared to the first stage, the cross country variation is less pronounced: estimates for most countries are between -0.6 and -0.4; moreover, the country patterns seen so far do not emerge as clearly in this case: for example, the propensity of Italian firms is only marginally lower than that of the Finnish or Swedish ones, and higher that the Austrians'.

In Model 2 we insert the firms' controls: the log of employment and of the operating revenues over total asset and the share of intangible assets. The coefficient on size is positive but substantially smaller than 1 (0.16 with a s.e. of 0.04), indicating that, conditional on having at least one patent, the propensity to patent decreases with firm size; more profitable firms have a higher PP while the coefficient of the share of intangibles is not significantly different from zero. Model 3 also includes a quadratic in size. Interestingly, also in this case we obtain a non monotonic effect of size: the linear term is positive and the quadratic negative. The minimum is reached at 48 employees, which is the 32^{th} percentile of the size distribution. Again, the results are very similar when we use total assets as an alternative measure of firm size. In nearly all the specifications the Mills ratio is significant, indicating that selection is indeed an issue.

In terms of the country dummies, it turns out that adding the firm characteristics significantly reduces their impact: the estimates are all substantially smaller in absolute values and, when including employment, not significant anymore for Belgium, Finland and Sweden; the sum of the coefficient drops (in absolute value) from -7.5 to -4.4 in Model 2. The estimates are further reduced when we include the quadratic in size: the cumulative value drops to -2.7, one third of the specification without firm characteristics, and the coefficient is not significant also for Austria. Figure 8, second panel reports the estimates of the country dummies for Model 1 (no firm characteristics) and 3 (quadratic in size): clearly, the drop is much more substantial than in the case of the probability of having at least one patent. This evidence therefore indicates that, conditional on patenting, country differences are substantially reduced when controlling for firm characteristics.

Table 9 reports the results for the second stage when the dependent variable is the ratio of patents over thousands employees. When we include no controls (Model 1), contrary to the other estimates, the country dummies are mostly positive and only statistically significant in three cases. This indicates that, for patenting firms, once we normalize the number of patents in terms of firm employment, the cross country differences basically disappear. When controlling for size, signs switch for some countries, but remain mostly statistically insignificant. This is the case throughout the board: this confirms that patenting firms are fairly similar when comparing their PP. This conclusion is also apparent from the last panel of Figure 8, which shows small and variable estimates for country effects.

4.3 Discussion

The analysis of the firm data offers some interesting insights on patenting activity. First of all, in general, the role of country does become less important after controlling for firm characteristics. This means that cross country differences are not due simply to the fact that some countries have an overall more favorable environment for innovation, but also they have different industrial structures.

Second, if we distinguish between engaging in patent activity at all and PP conditional on being engaged, the country effect is more relevant for the first stage: in some countries firms are more likely to engage in innovative activity, even after controlling for firm characteristics. Instead, conditional on engaging, the country effect is less relevant to explain the amount of patent activity that occurs at the level of the firm; moreover, the differences become substantially smaller when controlling for firm characteristics. These findings suggest that there is some sort of threshold that determines if a firm engages in patent activity and that this threshold is different across countries. This is an important determinant of differences in patenting activity. Given that large firms are likely to patent in any country, it indicates that small and medium sized firms are less likely to patent in laggard countries: this is particularly important for Southern European countries, that exhibit a higher share of small firms (Pagano and Schivardi 2003).

Once the firm has passed the patenting threshold, the nationality seems to be less relevant to determine the amount of innovation produced. One interpretation is that patenting firms are more internationally oriented and less sensitive to home country effects. Stated differently, after controlling for firms' characteristics, firms are fairly similar across countries in terms of patenting behavior.

Third, as far as the firms' characteristics are concerned, our findings contribute to the classic debate on firm size and innovation (Schumpeter 1942).[8] It is well-known that firm size increases the probability that a firm engages in innovation activity (Bound et al, 1984; Cohen et al, 1987). This fact is confirmed by our results. Conditional on innovating, the role of size is more complex. When we only allow for a linear relation between size and patents, the elasticity is below 1 and, correspondingly, PP decreases with size. This is in line with what found by Bound et al (1984) and Acs and Audretsch (1990), according to which the amount of patents per R&D expenditure declines with size. When we allow for a nonlinear relation between size and patenting activity, we find it to be non monotone: the propensity first declines and then increases with size. One possible interpretation is that, while large firms do enjoy advantages of scale and scope in innovating, there is also an important role for small, young, innovative firms that bring new products to the markets.

The policy implications of our findings are twofold. First, countries with a size structure with a prevalence of small firms might suffer from the fact that they are less likely to engage in patenting activity. Policies that induce such firms to grow would then increase their innovation output. At the same time, given that we find a less clear cut relation between PP and size for patenting firms, it is not clear that increasing further the size of large firms would result in an increase in patents.[9] Second, the evidence suggests that, even after controlling for firms' characteristics, the probability of engaging in patenting activity differs across countries. This indicates that policies that help firms to pass the patenting threshold, for example through tax incentives, grants or technical assistance for newly patenting firms, could have a large impact on country level patent activity. Moreover, given that patents area a mean of recovering R&D expenditures by granting temporary monopoly profits, increasing the PP would potentially increase the returns to innovation and therefore the incentives to undertake R&D activities.

[8]See Scherer (1991) and Cohen and Klepper (1991) for an overview of this debate.

[9]Taken literally, our results with the quadratic in size for patenting firms would suggest that a policy that aims at maximizing patents should reduce the number of medium size firms, in the range of 20-40 employees.

References

Acs, Z. S., and D. B. Audretsch (1990) *Innovation and Small Firms* (Cambridge, MA: MIT Press)

Audretsch, D.B., and R.P. Feldman (2004) 'Knowledge Spillovers and the Geography of Innovation.' In *Handbook of Regional and Urban Economics, vol. 4,* ed. Henderson V. and J.F. Thisse (Elsevier)

Bartelsman, E., S. Scarpetta, and F. Schivardi (2005) 'Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries.' *Industrial and Corporate Change* 14, 365–391

Bound, J., C. Cummins, Z. Griliches, B. Hall, and A. Jaffe (1984) 'Who does R&D and Who Patents?' In *R&D, Patents and Productivity,* ed. Z. Griliches (Chicago: University of Chicago Press for NBER)

Cohen, W., and S. Klepper (1991) 'A reprise of size and R&D.' *The Economic Journal* 106, 925–951

Cohen, W., R. Levin, and D. Mowery (1987) 'Firm Size and R&D Intensity.' *Journal of Industrial Economics* 35, 543–565

Crepon, B., E. Duguet, and J. Mairesse (1998) 'Research, Innovationa and Productivity: an Econometric Analysis at the Firm Level.' *Economics of Innovation and New Technology* 7(2), 115–158

Furman, J., M. Porter, and S. Stern (2002) 'The Determinants of National Innovative Capacity.' *Research Policy* 31, 899–933

Griliches, Z. (1990) 'Patent Statistics as Economic Indicators.' *Journal of Economic Literature* 28, 1661–1707

Hall, B. H. (2004) 'Exploring the Patent Explosion.' *NBER Working Papers.* n. 10605

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001) 'The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools.' *NBER Working Papers.* n. 8498

Heckman, J. (1979) 'Sample Bias as a Specification Error.' *Econometrica* (1), 153–161

Johnson, D. K. N. (2002 (5)) 'The OECD Technology Concordance (OTC): Patents by Industry of Manufacture and Sector of Use.' *STI Review*

Kortum, S., and J. Lerner (1998) 'Stronger Protection or Technological Revolution: What Is Behind the Recent Surge in Patenting.' *Carnegie-Rochester Conference Series on Public Policy* (48), 247–304

Paci, R., and F. Pigliaru (2002) 'Technological Diffusion, Spatial Spillovers and Regional Convergence in Europe.' Forthcoming in The European Monetary Union and Regional Convergence. Springer

Pagano, P., and F. Schivardi (2003) 'Firm Size Distribution and Growth.' *The Scandinavian Journal of Economics* 105, 255–274

Pakes, A., and Z. Griliches (1984) 'Patents and R&D at the Firm Level: a First Look.' In *R&D, Patents and Productivity,* ed. Z. Griliches (Chicago: University of Chicago Press for NBER)

Pavitt, K. (1984) 'Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory.' *Research Policy* 13, 343–373

Scherer, F. (1991) 'Changing Perspectives on the Firm Size Problem.' In *Innovation and Technological Change: An International Comparison,* ed. Z. Acs and D. Audretsch (New York: Harvester Wheatsheaf)

Schmookler, J. (1972) *Patents, Innovation and Economic Change: Data and Selected Essays* (Harvard University Press)

Schumpeter, J. (1942) *Capitalism, Socialism and Democracy* (Harper and Row)

Verspagen B., T. van Moergastel, and M. Slabbers (1994) 'MERIT concordance table: IPC - ISIC (rev. 2).' *MERIT Research Memorandum*

Table 1: *Patent share, by country and year, EU-15 countries.*

Year	AT	BE	DE	DK	ES	FI	FR	GB	GR	IE	IT	LU	NL	PT	SE	Total
1978	2.14	2.11	46.22	0.88	0.29	0.26	18.25	18.33	0.02	0.08	2.19	0.39	4.82	0.02	4.00	100
1979	2.41	1.95	44.71	0.92	0.22	0.34	20.23	15.79	0.01	0.24	3.05	0.33	5.73	0.01	4.06	100
1980	2.03	1.54	44.74	1.12	0.30	0.72	18.65	15.92	0.04	0.19	4.51	0.23	5.60	0.00	4.41	100
1981	1.96	1.25	42.17	1.08	0.26	0.59	19.18	16.80	0.06	0.11	4.94	0.38	6.54	0.01	4.68	100
1982	2.18	1.58	41.15	1.05	0.39	0.59	19.27	15.96	0.03	0.18	5.07	0.34	6.99	0.01	5.23	100
1983	2.10	1.26	42.38	1.13	0.51	0.69	17.76	16.04	0.04	0.16	5.77	0.33	7.08	0.01	4.72	100
1984	2.55	1.56	42.79	1.04	0.41	0.82	17.22	15.40	0.05	0.18	5.90	0.34	6.94	0.01	4.80	100
1985	2.41	1.36	42.25	1.07	0.48	0.83	17.87	15.58	0.02	0.12	6.59	0.38	6.61	0.02	4.41	100
1986	2.03	1.49	41.81	0.99	0.53	0.98	17.61	15.99	0.05	0.21	7.14	0.42	6.78	0.02	3.94	100
1987	2.26	1.60	42.56	0.95	0.57	1.13	17.92	14.36	0.04	0.26	7.43	0.30	6.83	0.03	3.76	100
1988	2.20	1.40	42.00	1.08	0.56	1.24	18.25	15.25	0.06	0.18	7.05	0.26	7.11	0.02	3.33	100
1989	2.20	1.54	40.81	1.29	0.73	1.39	18.87	14.45	0.08	0.24	7.62	0.17	7.23	0.01	3.36	100
1990	2.37	1.35	40.68	1.31	0.77	1.60	19.21	13.84	0.08	0.28	7.74	0.18	6.88	0.02	3.69	100
1991	2.27	1.70	40.99	1.41	0.92	1.55	19.01	13.29	0.08	0.24	7.83	0.23	6.75	0.02	3.71	100
1992	2.09	1.74	41.47	1.49	0.87	1.97	18.22	13.99	0.10	0.32	6.92	0.21	6.48	0.05	4.09	100
1993	2.15	1.94	40.56	1.56	1.00	2.18	18.82	12.90	0.07	0.36	7.29	0.40	6.57	0.07	4.16	100
1994	2.07	1.85	42.12	1.64	0.99	2.58	17.40	12.45	0.07	0.28	6.28	0.20	7.08	0.06	4.94	100
1995	1.88	1.85	42.37	1.62	1.00	2.44	16.81	11.95	0.08	0.46	6.62	0.19	7.37	0.04	5.33	100
1996	1.85	1.89	44.28	1.70	0.96	2.64	15.98	10.99	0.08	0.42	6.45	0.27	6.80	0.04	5.64	100
1997	2.10	2.05	44.94	1.54	1.12	2.91	15.80	10.16	0.13	0.43	6.42	0.26	6.49	0.05	5.61	100
1998	1.85	1.80	45.62	1.66	1.04	3.00	15.53	10.43	0.14	0.46	6.05	0.29	6.61	0.06	5.45	100
1999	1.74	1.66	44.81	1.56	1.11	3.38	15.62	10.64	0.10	0.47	6.20	0.26	7.03	0.05	5.37	100
2000	1.96	1.56	45.04	1.65	1.09	3.25	15.14	10.52	0.09	0.41	6.36	0.22	7.51	0.06	5.12	100
2001	2.34	1.49	49.09	1.05	1.13	2.77	16.15	8.06	0.12	0.41	7.84	0.16	5.95	0.05	3.39	100
Average	2.13	1.65	43.15	1.28	0.72	1.66	17.70	13.71	0.07	0.28	6.22	0.28	6.66	0.03	4.47	-

Source: authors' calculations using EPO data.

Table 2: *Decomposition of PP*

Country	Total	Efficiency	Intensity	Cross term
AT	-18.7	1.2	-19.7	-0.2
BE	-36.0	-49.1	25.7	-12.6
DE	94.8	51.4	28.6	14.7
DK	-0.8	-27.1	36.1	-9.8
ES	-88.4	-82.7	-32.5	26.9
FI	136.5	3.6	128.2	4.7
FR	4.4	-22.4	34.4	-7.7
GB	-40.1	-10.5	-33.1	3.5
GR	-96.1	-94.4	-30.3	28.6
IE	-59.6	-44.8	-26.9	12.0
IT	-53.9	-28.0	-36.0	10.1
LU	38.7	-3.5	43.6	-1.5
NL	54.8	39.5	11.0	4.3
PT	-97.9	-95.2	-56.1	53.4
SE	101.2	23.8	62.5	14.9

Note: "Total" is the percentage difference in PP with respect to the EU15 average, $\frac{p_j - \bar{p}}{\bar{p}}$. The three terms of the decomposition are defined in equation (1).

Table 3: *Sectoral patent propensity: patent per 1000 employees*

	AT	BE	DE	DK	ES	FI	FR	GB	IE	IT	NL	PT	SE	EU-15	Coeff. var.
Food, beverages and tob.	0.42	1.27	1.22	2.05	0.18	1.94	0.62	0.87	0.28	0.52	4.79	0.01	1.28	0.17	1.1
Textiles and leather	1.38	1.55	3.29	2.64	0.11	1.19	1.45	0.78	0.50	0.63	2.89	0.00	7.07	0.17	1.0
Wood products	0.95	0.86	1.85	0.88	0.17	0.41	1.07	0.97	0.28	0.86	1.99	0.03	0.84	0.29	0.7
Paper and publishing	1.48	1.73	1.51	0.52	0.10	1.66	0.99	0.56	0.26	0.58	2.31	0.04	1.31	0.34	0.7
Chemical (excl. Pharm.)	1.09	0.56	1.51	1.40	0.10	1.09	0.94	0.92	0.28	0.44	1.54	0.03	1.04	2.83	0.6
Pharmaceutical	0.62	0.86	1.72	1.56	0.20	0.87	0.75	1.22	0.26	0.31	1.72	0.09	1.48	7.61	0.7
Petroleum and plastic	0.71	0.50	0.92	1.22	0.19	1.31	1.92	0.65	2.78	0.70	2.97	0.00	0.73	0.23	0.8
Other non-metallic mineral pr.	1.41	0.65	1.70	1.26	0.12	1.01	1.53	0.70	0.32	0.56	1.50	0.02	2.03	0.69	0.6
Basic metals (excl. non ferrous)	2.09	0.64	1.86	0.40	0.04	1.50	0.64	0.44	0.00	0.52	0.82	0.00	1.58	0.47	0.9
Metal products (excl. machin.)	1.47	0.59	1.76	0.98	0.15	1.60	0.92	0.68	0.54	0.57	1.32	0.06	1.52	1.33	0.6
Machinery n.e.c.	0.89	0.95	1.44	0.83	0.18	1.43	1.08	0.57	0.75	0.55	1.48	0.03	1.22	2.18	0.5
Office machinery and computers	1.82	2.38	1.33	1.57	0.08	10.47	0.84	0.60	0.12	0.39	4.03	0.00	3.92	9.69	1.3
Electrical machinery n.e.c.	1.22	0.26	1.26	0.72	0.17	1.98	1.46	0.54	0.17	0.44	5.06	0.00	1.23	2.52	1.2
Communication equipment	0.17	0.31	1.60	0.38	0.10	2.29	0.90	0.55	0.12	0.18	2.85	0.02	1.63	6.35	1.1
Motor vehicles	1.08	0.17	1.33	0.00	0.12	2.34	0.95	0.49	8.07	0.87	2.08	0.00	0.81	2.24	1.5
Other transport equipment	2.01	0.36	1.95	0.69	0.19	1.13	0.98	0.29	0.44	0.50	0.67	0.08	1.42	0.86	0.8
Precision instruments	0.90	1.96	1.19	0.95	0.28	1.94	0.85	0.82	0.31	0.42	2.32	0.05	2.42	6.94	0.7
Manufacturing n.e.c.	0.77	1.03	1.67	0.70	0.19	2.38	1.36	0.77	0.00	0.53	2.42	0.03	1.77	1.13	0.8
All manufacturing	0.85	0.65	1.61	0.94	0.12	2.05	1.04	0.71	0.45	0.42	2.32	0.02	1.80	1.58	0.7
Coeff. of variation	0.46	0.68	0.31	0.61	0.38	1.07	0.32	0.32	2.21	0.32	0.52	1.01	0.81	1.17	

All entries are expressed as ratio to the value for the total EU-15 in the last columns. Source: sectoral employment data are from the Eurostat Newcronos Database; patent data from the EPO database.

Table 4: *Patent Propensity decomposition: sectoral propensity and composition.*

Country	Total	Sectoral propensity	Sectoral	Interaction
AT	-14.5	-1.3	-1.7	-11.5
BE	-34.8	-10.3	-7.6	-16.9
DE	61.3	50.4	10.3	0.5
DK	-5.7	-2.4	-1.1	-2.3
ES	-88.1	-83.5	-25.4	20.8
FI	104.8	107.6	8.8	-11.6
FR	3.7	0.1	6.6	-3.1
GB	-29.1	-30.2	2.0	-0.9
IE	-58.3	-17.0	68.4	-109.8
IT	-57.6	-53.7	-9.9	6.0
NL	131.9	130.2	3.1	-1.4
PT	-98.2	-96.5	-43.6	41.9
SE	80.2	66.4	16.1	-2.2

Note: "Total" is the percentage difference in PP with respect to the EU15 average, $\frac{p_j - \bar{p}}{\bar{p}}$. The three terms of the decomposition are defined in equation (2).

Table 5: *AMADEUS-population comparisons*

Country	No. Firms		Employment		Avg size		Share manuf. emp.		Pantent Prop		No. Patents	
	Amad.	Amad/Pop. (%)	Amad.	Amad/Pop. (%)	Amad.	Pop.	Amad.	Pop.	Amad.	Pop.	Amad.	Amad/Pop. (%)
AT	159	0.1	119694	5.9	753	10.0	0.33	0.30	0.38	0.42	160	18.80
BE	6447	2.8	970524	47.9	151	8.7	0.39	0.31	0.26	0.33	258	38.17
DE	1554	0.1	1744727	9.3	1123	11.6	0.53	0.40	2.50	1.04	8144	41.74
DK	2533	1.3	513931	32.0	203	8.1	0.37	0.30	0.59	0.44	321	44.96
ES	9243	0.4	2269599	25.2	246	4.1	0.31	0.27	0.06	0.05	144	30.38
FI	2105	1.2	449078	39.3	213	6.2	0.43	0.38	2.00	1.23	964	68.42
FR	14439	0.7	4505429	34.8	312	6.7	0.37	0.30	0.53	0.51	3239	49.38
GB	13578	0.9	6041666	38.6	445	10.6	0.27	0.25	0.15	0.29	1256	27.55
GR	2150		424736		198		0.37		0.002		1	2.50
IE	517	0.8	101545	14.2	196	11.2	0.33	0.36	0.03	0.25	5	2.82
IT	13910	0.4	2523935	28.1	181	2.5	0.43	0.45	0.30	0.31	906	32.89
LU	58		19498		336		0.39		0.67		13	13.68
NL	1673	0.4	278714	7.1	167	9.7	0.35	0.22	0.86	0.83	624	19.19
PT	506	0.1	221807	9.4	438	4.6	0.31	0.38	0.02	0.01	8	29.63
SE	5606	1.3	957792	41.7	171	5.2	0.40	0.33	0.76	0.97	763	34.40
EU-15	74478	0.6	21142675	25.3	284	6.3	0.36	0.32	0.52	0.52	16806	38.80

The population data are from OECD. For Ireland they exclude the Energy and the Construction sectors and for the Netherlands the Construction one.

Table 6: *Descriptive statistics, AMADEUS regression sample*

	No. obs.	Empl.		Tot. Assets		Op. rev/empl.		Intang./fixed assets		No. patents.		Pat./th.emp	
		Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.
Non patenting firms													
AT	84	1008.0	3850.9	181.0	408.9	268.3	272.6	0.8	3.4				
BE	4,857	154.6	841.5	92.1	526.9	1848.3	6256.9	52.5	150.0				
DE	843	563.5	1018.5	323.3	1298.9	1321.4	6036.2	17.6	75.0				
DK	1,472	219.5	516.1	76.7	513.6	631.6	1345.1	88.6	182.6				
ES	6,737	233.3	934.4	77.6	625.6	657.2	2035.5	132.0	204.8				
FI	1,528	187.6	658.7	57.4	225.6	824.7	3148.3	104.4	182.5				
FR	10,800	250.3	897.3	97.0	1120.1	727.1	2516.2	130.5	216.4				
GB	8,442	351.5	948.1	120.2	1185.4	840.9	3388.0	51.5	167.9				
GR	1,756	170.7	902.4	44.6	214.3	475.0	1170.6	100.7	166.6				
IE	184	174.1	298.4	56.8	138.8	860.6	2378.9	41.9	150.5				
IT	9,692	148.7	618.9	52.6	427.4	1101.9	3227.4	129.4	197.6				
LU	36	377.8	526.6	141.6	436.0	413.1	698.5	18.8	53.7				
NL	807	188.5	378.2	115.8	383.7	1755.0	6569.4	40.5	149.3				
PT	418	424.0	1380.0	120.7	494.9	595.3	2477.4	53.7	138.5				
SE	3,728	165.9	509.9	74.9	451.9	597.3	1104.9	47.2	149.8				
EU15	51,384	231.5	839.5	88.5	816.0	919.8	3374.3	96.6	190.2				
Patenting firms													
AT	8	423.4	261.7	124.7	93.4	419.4	324.9	5.5	10.4	1.6	1.8	9.1	15.0
BE	40	841.9	1291.0	497.5	906.0	385.9	317.9	123.4	192.9	6.2	12.7	25.4	78.5
DE	131	5827.2	22453.3	2529.8	12018.5	317.7	420.7	20.4	69.0	31.6	125.0	11.7	42.5
DK	58	879.8	1438.7	343.6	745.1	342.0	850.0	88.8	178.9	4.8	8.9	12.5	18.4
ES	68	666.8	1158.1	905.7	5812.4	306.2	493.4	164.2	218.7	2.0	2.4	24.2	121.5
FI	64	740.8	1705.2	797.4	2208.0	263.2	282.1	129.6	206.7	14.0	79.2	37.3	97.0
FR	294	2053.7	10576.9	1655.0	10782.0	465.5	2162.9	101.1	179.5	6.5	22.6	27.0	104.9
GB	227	1274.4	2507.2	632.5	2387.2	253.5	283.6	91.5	211.5	3.8	7.6	15.9	50.6
GR	1	1000.0		105.1		82.0		10.8		1.0		1.0	
IE	2	170.0	113.1	16.5	5.6	269.0	75.0	0.0	0.0	1.5	0.7	13.1	12.9
IT	233	799.6	2404.9	533.2	2828.8	456.8	1604.8	125.2	173.5	2.9	4.7	26.6	88.7
LU	3	415.0	126.7	911.4	1321.2	893.7	1176.0	11.2	19.3	4.3	5.8	8.6	9.8
NL	15	1219.5	2493.5	297.1	628.4	215.7	124.8	21.0	30.1	14.5	37.5	75.0	230.5
PT	2	401.0	101.8	25.9	2.6	76.0	9.9	21.9	13.4	2.5	2.1	7.1	7.1
SE	122	861.7	2198.1	598.3	1820.3	307.7	387.2	52.3	147.8	5.9	18.7	50.3	164.3
EU15	1,268	1695.9	9104.4	1040.7	6859.5	365.0	1289.1	93.6	179.4	8.0	46.7	25.6	96.9

Table 7: *First stage: Probit model*

	Model 1		Model 2		Model 3		Model 4		Model 5	
	dF/dx*100	S.E.	dF/dx*100	S.E.	dF/dx*100	S.E.	dF/dx*100	S.E.	dF/dx*100	S.E.
$Dpat99$	*16.378*	0.009	*12.760*	0.008	*13.048*	0.008	*11.356*	0.008	*11.651*	0.008
AT	-0.225	0.001	-0.141	0.001	-0.130	0.002	-0.097	0.002	-0.095	0.002
BE	*-0.428*	0.000	*-0.303*	0.000	*-0.319*	0.000	-0.286	0.000	*-0.301*	0.000
DK	*-0.293*	0.001	-0.190	0.001	-0.188	0.001	-0.166	0.001	-0.174	0.001
ES	*-0.458*	0.000	*-0.334*	0.000	*-0.346*	0.000	*-0.308*	0.000	*-0.321*	0.000
FI	*-0.306*	0.000	-0.170	0.001	*-0.173*	0.001	*-0.158*	0.001	*-0.172*	0.001
FR	*-0.411*	0.001	*-0.286*	0.001	*-0.293*	0.001	*-0.250*	0.001	*-0.264*	0.001
GB	*-0.443*	0.001	*-0.333*	0.000	*-0.344*	0.000	*-0.291*	0.000	*-0.304*	0.000
GR	*-0.453*	0.000	*-0.365*	0.000	*-0.385*	0.000	*-0.339*	0.000	*-0.352*	0.000
IE	*-0.369*	0.001	-0.276	0.001	-0.284	0.001	-0.251	0.001	-0.262	0.001
IT	*-0.490*	0.001	*-0.322*	0.001	*-0.334*	0.001	*-0.301*	0.000	*-0.315*	0.000
LU	-0.193	0.003	-0.109	0.003	-0.074	0.003	-0.138	0.002	-0.147	0.002
NL	*-0.386*	0.000	*-0.300*	0.000	*-0.311*	0.000	*-0.272*	0.000	*-0.284*	0.000
PT	*-0.365*	0.001	*-0.291*	0.001	*-0.307*	0.001	*-0.269*	0.001	*-0.279*	0.001
SE	*-0.334*	0.000	*-0.178*	0.001	*-0.185*	0.001	*-0.168*	0.001	*-0.182*	0.001
$LnEmpl$			*0.113*	0.000	*-0.139*	0.001				
$LnEmpl^2$					*0.026*	0.000				
$LnAssets$							*0.177*	0.000	-0.035	0.001
$LnAssets^2$									*0.009*	0.000
$LnOp.Rev.Turn$			*0.063*	0.000	0.037	0.000				
$LnOp.Rev.Empl$							*-0.070*	0.000	*-0.073*	0.000
$Sh.Intangibles$			*0.253*	0.001	*0.273*	0.001	*0.269*	0.001	*0.280*	0.001
N. obs.	52652		52652		52652		52542		52542	
LR χ^2	*5264.29*		*5444.92*		*5473.3*		*5507.72*		*5512.57*	
$Ps.R^2$	0.4403		0.4555		0.4578		0.4618		0.4622	

Dependent variable: a dummy equal to 1 if the firm applies for at least 1 patent in year 2000. All regressions include industry dummies. The excluded country is Germany. Dpat99 is a dummy equal to 1 if the firm had applied for at least 1 patent before 2000. Coefficients represent the marginal effects. Values in italics are significant at 5%.

Table 8: *Second stage: Patent count*

	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
AT	*-1.083*	0.386	*-0.797*	0.350	-0.512	0.339	*-0.664*	0.344	-0.505	0.340
BE	*-0.419*	0.193	-0.196	0.176	-0.113	0.170	-0.213	0.173	-0.252	0.171
DK	*-0.534*	0.167	*-0.343*	0.152	*-0.255*	0.147	*-0.380*	0.149	*-0.383*	0.147
ES	*-0.818*	0.163	*-0.553*	0.150	*-0.410*	0.145	*-0.575*	0.146	*-0.555*	0.144
FI	*-0.539*	0.162	-0.191	0.149	-0.125	0.144	*-0.324*	0.145	*-0.379*	0.143
FR	*-0.479*	0.114	*-0.301*	0.104	*-0.224*	0.101	*-0.309*	0.102	*-0.325*	0.101
GB	*-0.592*	0.120	*-0.435*	0.109	*-0.349*	0.105	*-0.416*	0.107	*-0.426*	0.105
GR	0.045	1.060	-0.383	0.975	-0.350	0.940	-0.521	0.958	-0.361	0.942
IE	-1.015	0.740	-0.541	0.675	-0.397	0.651	-0.331	0.662	-0.469	0.652
IT	*-0.626*	0.120	*-0.331*	0.111	*-0.263*	0.107	*-0.385*	0.108	*-0.402*	0.106
LU	-0.935	0.620	-0.553	0.561	-0.170	0.543	-0.665	0.550	-0.451	0.543
NL	-0.032	0.286	0.199	0.260	0.274	0.251	0.148	0.255	0.148	0.251
PT	0.008	0.743	0.179	0.680	0.317	0.656	0.272	0.668	0.261	0.657
SE	*-0.529*	0.135	-0.195	0.124	-0.136	0.120	*-0.284*	0.121	*-0.304*	0.119
LnEmpl			*0.154*	0.035	*-0.434*	0.072				
$LnEmpl^2$					*0.056*	0.006				
LnAssets							*0.282*	0.020	*-0.625*	0.144
$LnAssets^2$									*0.037*	0.006
LnOp.Rev.Turn			*0.120*	0.032	0.052	0.032				
LnOp.Rev.Empl							-0.062	0.037	-0.066	0.036
Sh.Intangibles			-0.123	0.155	-0.073	0.149	-0.144	0.152	-0.145	0.149
Const	*1.822*	0.321	*-1.121*	0.408	*1.065*	0.452	*-1.921*	0.434	*3.613*	0.965
λ	*-0.395*	0.052	*-0.216*	0.051	*-0.229*	0.049	*-0.175*	0.052	*-0.218*	0.052
N. obs.	52652		52652		52652		52542		52542	
N. nonzero obs.	1268		1268		1268		1265		1265	
χ^2	*537.01*		*860.42*		*987.820*		*953.790*		*1005.290*	
ρ	-0.36681		-0.227		-0.248		-0.18811		-0.236	

Dependent variable: log of number of patent applications in year 2000. All regressions include industry dummies. The excluded country is Germany. λ is the Mills ratio from the first stage regression.

Table 9: *Second stage: Patents per 1000 employees*

	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
AT	-21.611	34.449	*-67.413*	30.279	-11.281	24.713	-52.111	31.719	-38.840	31.391
BE	10.435	17.284	*-25.264*	15.254	-7.518	12.413	-7.064	15.962	-10.081	15.768
DK	-0.643	14.906	-19.909	13.112	-1.663	10.691	-5.519	13.704	-5.593	13.527
ES	14.072	14.580	-13.262	12.921	16.352	10.554	5.279	13.507	7.212	13.337
FI	*31.303*	14.423	-5.415	12.850	7.967	10.474	*22.273*	13.333	17.740	13.187
FR	14.229	10.135	-9.788	8.980	6.475	7.332	1.634	9.377	0.500	9.261
GB	5.315	10.668	-10.519	9.395	7.806	7.671	-1.059	9.838	-1.707	9.712
GR	1.581	96.744	49.392	84.792	64.022	68.638	36.291	88.884	50.517	87.743
IE	-1.080	66.563	-55.963	58.443	-27.755	47.470	-34.563	61.226	-45.979	60.460
IT	17.216	10.650	*-19.681*	9.535	-5.262	7.777	3.697	9.910	2.453	9.786
LU	2.911	55.118	-50.944	48.434	22.711	39.532	-20.562	50.637	-2.856	50.079
NL	*67.945*	25.611	*45.478*	22.484	*61.892*	18.287	*63.432*	23.522	*63.540*	23.218
PT	5.603	67.274	4.997	59.017	34.217	47.871	10.408	61.837	9.897	61.026
SE	*32.014*	12.049	-7.792	10.706	4.321	8.726	18.019	11.118	16.349	10.982
LnEmpl			*-47.035*	3.053	*-164.312*	5.239				
$LnEmpl^2$					*11.168*	0.447				
LnAssets							*-12.959*	1.890	*-89.898*	13.309
$LnAssets^2$									*3.154*	0.543
LnOp.Rev.Turn			*17.124*	2.762	3.253	2.302				
LnOp.Rev.Empl							*28.678*	3.405	*28.394*	3.361
Sh.Intangibles			*-37.201*	13.376	*-28.285*	10.880	-24.286	13.991	-24.432	13.810
Const	-9.617	28.774	*124.445*	35.300	*567.717*	32.901	5.054	40.103	*474.876*	89.200
λ	5.428567	4.672	-6.922	4.438	*-12.976*	3.598	0.212	4.806	-3.846	4.781
N. obs.	52652		52652		52652		52542		52542	
N. nonzero obs.	1268		1268		1268		1265		1265	
χ^2	*521.82*		*1029.81*		*2018.820*		*831.92*		*887.390*	
ρ	0.0585		-0.085		-0.194		0.00249		-0.046	

Dependent variable: number of patent applications over thousands employees in year 2000. All regressions include industry dummies. The excluded country is Germany. λ is the Mills ratio from the first stage regression.

Figure 1: *Trend in patent applications at the EPO.*

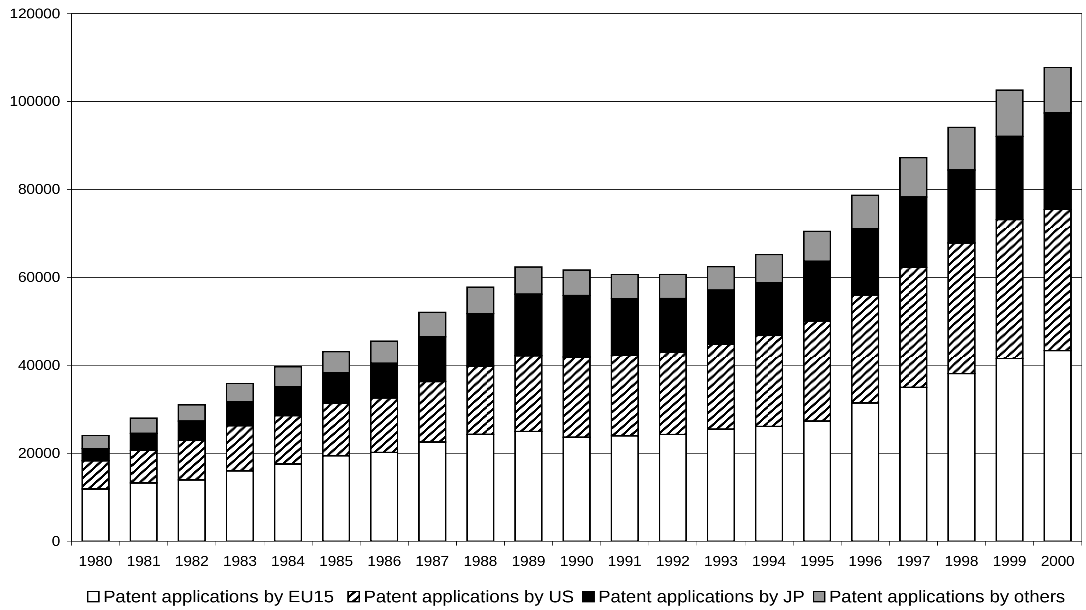
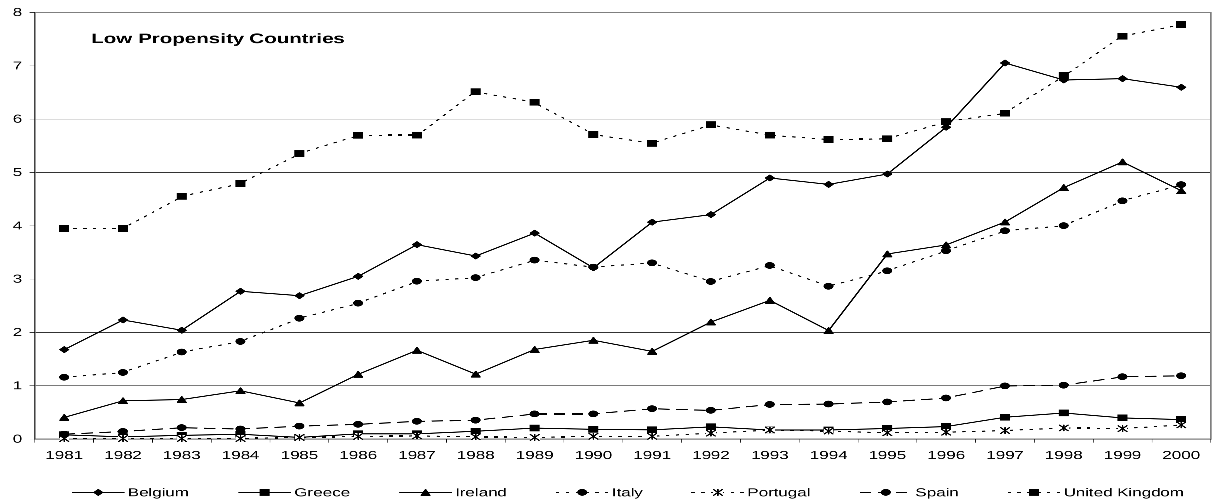


Figure 2: *Evolution of Patent Propensity (patens by 100.000 inhabitants.*



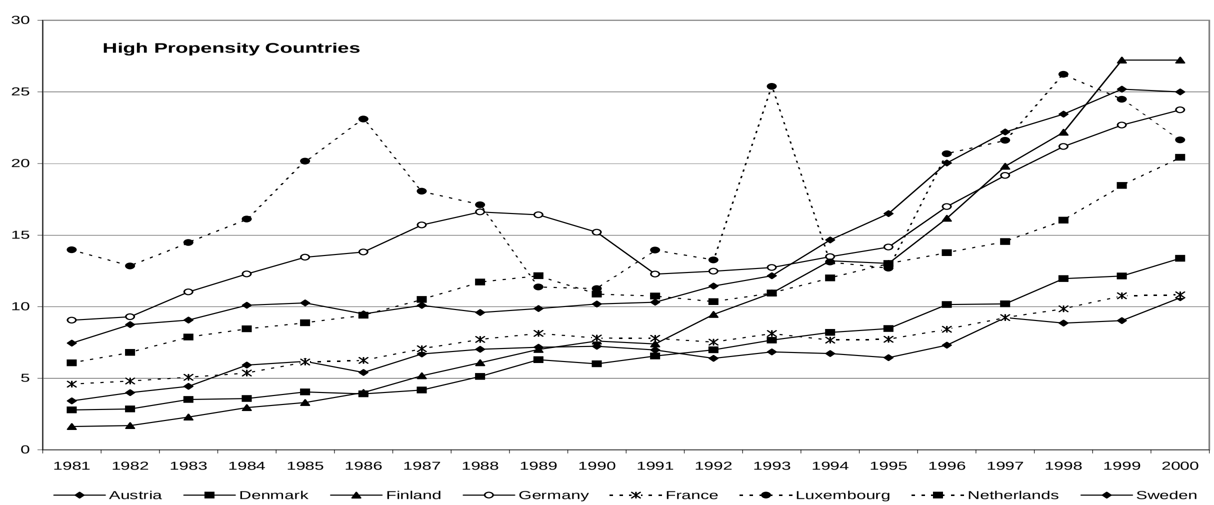


Figure 3: *Convergence in Patent Propensity.*

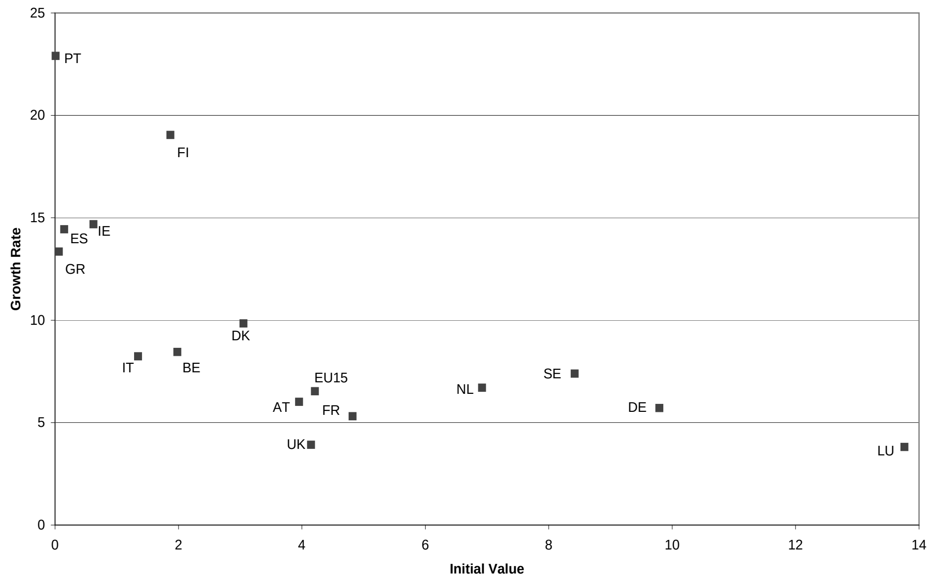
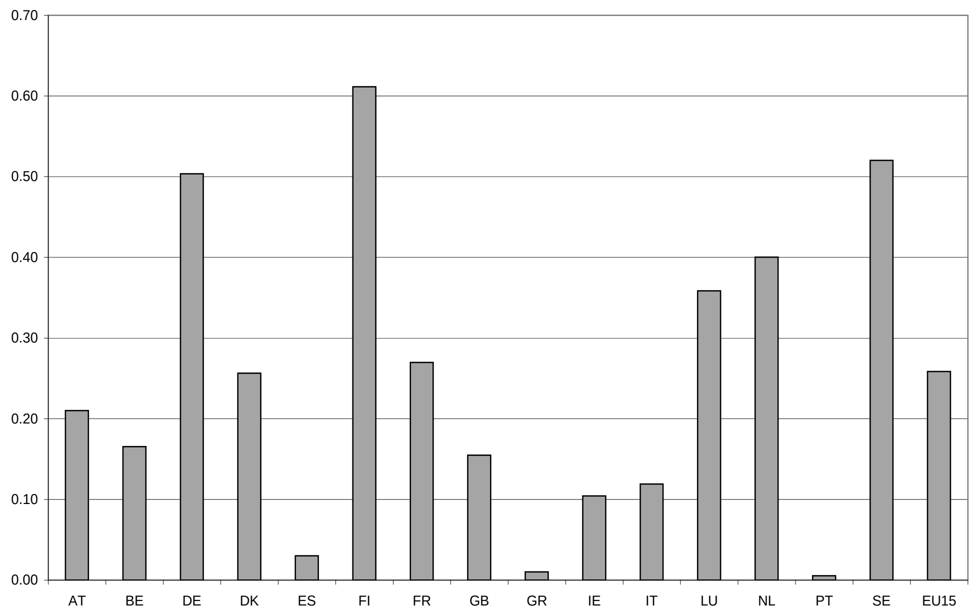
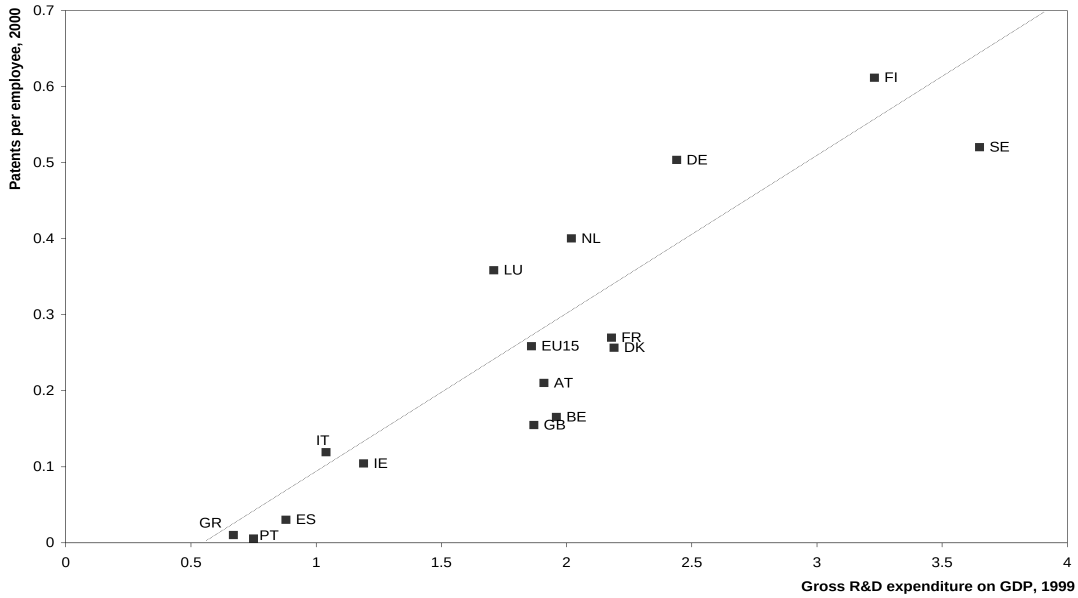


Figure 4: *Patents per thousands workers in 2000.*



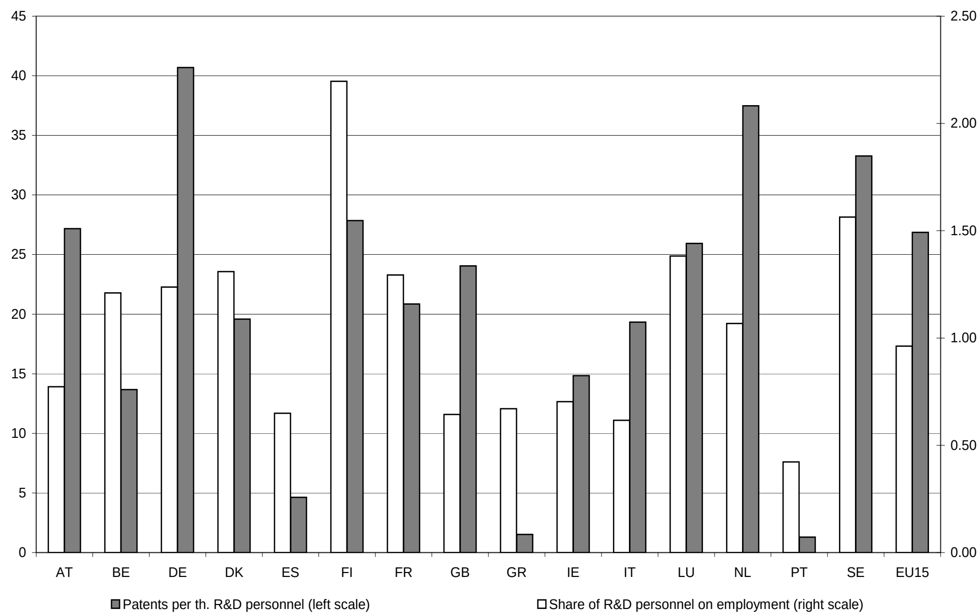
Sources: patents data are from the EPO; employment data from the OECD Main Science and Technology Indicators database.

Figure 5: *Patents and R&D intensity.*



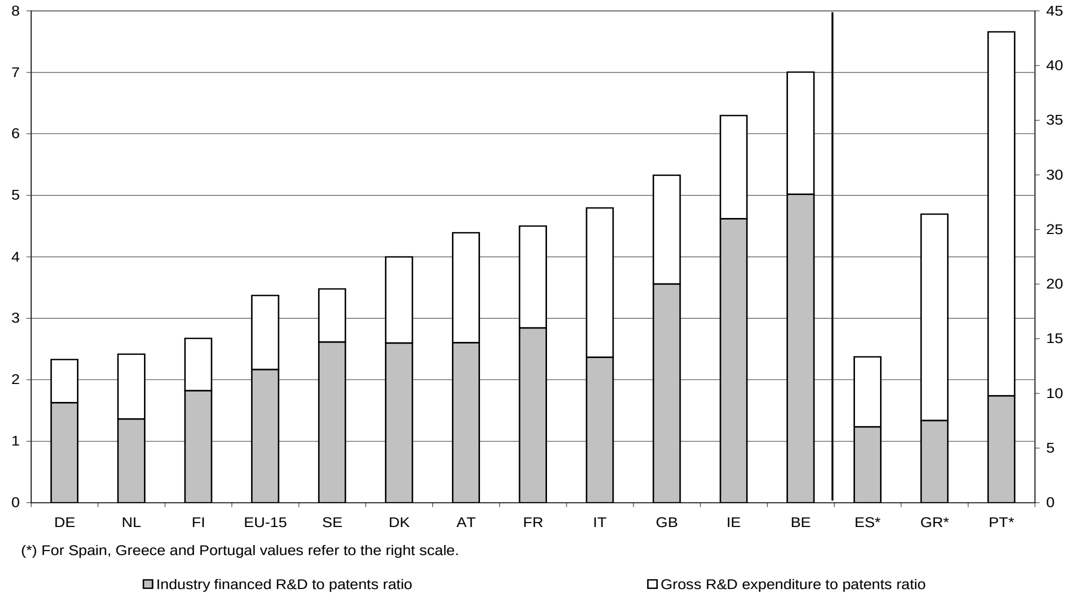
Sources: patents data are from the EPO; R&D to GDP ratio data from the OECD Main Science and Technology Indicators database.

Figure 6: *Patents and R&D personnel in 2000.*



Source: patents data are from the EPO; employment and R&D personnel data from the OECD Main Science and Technology Indicators database.

Figure 7: *Unit cost of a patent, in terms of R&D expenditure.*



Sources: patents data are from the EPO; R&D expenditures data from the OECD Main Science and Technology Indicators database.

Figure 8: *Country Dummies Estimates: probit estimates (first panel) and patent propensity (log of the number of patents in the second panel and patents per thousand employees in the third).*