Access to universities' public knowledge: Who's more nationalist?

Joaquín M. Azagra-Caro[♣]

Abstract Access to public knowledge is a prerequisite for the good functioning of developed economies. Universities strive and are also requested to contribute to this knowledge both locally and internationally. Traditional studies on the geography of knowledge flows have identified a localisation effect; however, these studies do not use the country as the unit of observation and hence do not explore national patterns. In this paper, we hypothesise that the localisation of university knowledge flows is directly related to share of firm expenditure on research and development. To test this hypothesis, we use references to universities in patent documents as indicators based on a data set of around 20,000 university references, for 37 countries in the period 1990-2007, resulting in panels of around 300-500 observations. We build indicators for the university knowledge flows both inside and outside the applicant country, which we explain as a function of some proxies for national size and research structure based on

[♣] J.M. Azagra-Caro (✉)
INGENIO (CSIC-UPV), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
Paper started at:
European Commission, Joint Research Centre (JRC)-Institute for Prospective Technological Studies (IPTS), Edificio Expo, C/Inca Garcilaso 3, E-41092 Sevilla, Spain; and
CSIC - Consejo Superior de Investigaciones Científicas, Institute of Public Goods and Policies (IPP-CCHS), C/Albasanz 26-28, E-28037 Madrid, Spain
e-mail: jazagra@ingenio.upv.es. Tel: (+34) 963877048 ext. 78439. Fax: (+34) 963877991

econometric estimations. We draw some conclusions as to the importance of national business scientific strength for fostering increased domestic university knowledge flows.

Keywords Universities; Knowledge flows; R&D expenditure

Introduction

Effective knowledge sharing is one of the ideals of the European Research Area (ERA), and includes access to the public knowledge base and, more specifically, to the knowledge bases of public universities. Providing access is complex since it should occur 'both within and across borders' (EC 2007: 16), i.e. on geographic levels. The challenge for the universities is clear: on the one hand, clusters of partnerships with universities should 'form and expand through virtual integration rather than geographical concentration' (EC 2007: 8); on the other hand, public-private cooperation allows universities to 'excel in addressing research and training needs at national and regional level' (EC 2007: 14).

Therefore, there is a need for policy to enable access to university knowledge inside and outside geographical borders. However, we lack the theoretical background required to understand national variations on both these dimensions. Our paper contributes by reviewing the academic literature on knowledge flows, measured through patent citations. The concept of 'knowledge flows' in this stream of literature can be reasonably compared to the concept of 'access to the public knowledge base' in the policy literature because both refer to links between current and past pieces of knowledge.

Knowledge flows measured through patent citations have been criticised for not distinguishing between examiner and applicant citations, because examiner citations may not express knowledge flows and may bias results (Alcácer and Gittelman 2006; Criscuolo and Verspagen 2008; Azagra-Caro et al. 2009). However, the concept of access to the public knowledge base does not require such a distinction: what matters is that a piece of knowledge is sufficiently widely known for anyone (patent applicant or examiner) to refer to it, i.e. it is accessible.

Since we focus on universities, this paper will also contribute to work on patent citation analysis in moving away from the usual distinction of type of literature (patent vs non-patent references, e.g. Callaert et al. 2006) to the institution type (university vs non-university references).

The measurement of knowledge flows inside and outside geographical borders is an emerging area of research. Some authors tackle only the outside dimension (Acosta et al. 2011); where both dimensions are addressed, it is usually for only a single country (Abramo et al. 2010). The third contribution of this paper is measurement of the phenomenon at macro level, for a large number of countries and years, and an empirical explanation for the observed variation.

Below we provide a review of the literature and our working hypothesis. We describe the data and methodology, present our results, discuss some limitations of this study and finish with some conclusions.

Literature review

Knowledge flows inside and outside the country: What patent citations tell us

It is accepted that citation linkages represent an indirect rather than a direct link between a cited paper and a citing patent (Meyer 2000) (science provides background knowledge and human capital), or an intertwined link (technology too can drive science). When studying a country's patents, patent citations are typically used to differentiate between national and foreign sources. For instance, the large number of non-Dutch-invented United States (US) Patent and Trademark Office (USPTO) patents citing Dutch-authored research papers indicates that 'the benefits of domestic basis science for technical inventions are not constrained by national barriers, but add to the global knowledge base to the benefit of all' (Tijssen 2001: 53). Among citations to foreign countries, it is possible to distinguish the particular country, i.e. Hu and Jaffe (2003) find that the Koreas' knowledge base relies more on Japanese than the US knowledge, whereas Taiwan's knowledge base draws on knowledge from both countries. In comparing the US and the EU, citations are an illustration of where the so-called European Paradox occurs, e.g. showing that US patents in biotechnology refer to EU papers more often than EU patents, but not in information technology (Verbeek et al. 2003). Citations are made by the patent examiner as well as the patent applicant, and it may be that each has a different preference for national or foreign citations. For instance, foreign applicants to the USPTO have the highest proportion of citations added by examiners (Alcácer et al. 2009).

All these examples illustrate that patent citations can be a good tool to measure access to the knowledge base inside and outside a country. However, they do not provide quantitative explanations about why each occurs.

Patent citations as an indication of the geography of knowledge flows: the localisation effect

A series of papers addressed the idea of a localisation effect in knowledge flows: patents tend to include more citations to prior art from the applicant's geographical country or region than from other countries or regions. However, most studies use the citation not the country or region as the unit of observation, and put it as a function of the characteristics of the patent to which the citation belongs.

Jaffe et al. (1993) is a seminal work in this field. Jaffe and colleagues estimate an econometric model to find the determinants of a variable equal to 1 if the applicant's country/state/city of the originating patent matches that of the citing patent, and 0 otherwise. Jaffe and colleagues consider it a function of a variable equal to 1 if the match is between the originating patent and a control patent (for the same technology and year of the citing patent), and 0 otherwise. They use USPTO data to find a positive relation, suggesting a higher probability of a domestic match in domestic citations.

Other studies reinforce this impression. Jaffe and Trajtenberg (1996), also based on USPTO data, explain the number of US citations as a function of the nationality of the citing applicant. Their finding that US patents show a higher propensity to include US citations than other geographic units (Canada, European Economic Community, Japan, and rest of the world) is described as a localisation effect. Jaffe and Trajtenberg (1999) generalise this localisation effect to other countries: patents originating in the United Kingdom (UK), France, Germany and Japan are cited more often in patent applications from applicants in those respective countries than from any other country.

Maurseth and Verspagen (2002) extend this research to introduce region as the unit of observation. They use European Patent Office (EPO) data to confirm previous findings for a sample of European countries: they regress the number of citations between two given regions on a measure of geographical distance between citing and cited patent. The sign of the estimated coefficient is always negative, which is evidence of localisation: more proximate regions are more likely to cite one another. This result holds even if we distinguish between examiner and applicant citations (Criscuolo and Verspagen 2008).

These studies tend to interpret the paper trail left by patent citations as knowledge spillovers. First, they show that there are geographic limits to these spillovers, but this is not the same as saying that location always matters for innovation because innovation depends also on the type of activity, stage in the industry life cycle and composition of activity within a location (Feldman 1999). Second, it is questionable whether these studies really capture knowledge spillovers because these flows could be interpreted as stemming from contractual arrangements, researcher mobility or other market mechanisms (Breschi and Lissoni 2001).

We should add that these studies also do not tackle the inside/outside dimension of knowledge flows referred to in the previous section. They also do not provide indications about in which countries knowledge flows will be more localised. However, they offer some insights that allow us to deduce testable hypotheses using country as the unit of observation. We demonstrate this in the following two sections.

Country as the unit of observation

In the studies referred to above, the unit of observation is rarely territorial (country or region); it is usually a patent citation. The exception is the work by Maurseth and Verspagen (2002), but their focus is on the distance between regions. Suppose we take country as the unit of observation in order to determine whether a citation is national. The relation between size and the share of national university references has not been hypothesised in previous works, but the literature about proximity has dealt with related issues, so we draw on it to build the following hypothesis:

Hypothesis 1. The larger the size of a country, the higher the share of references to the work of universities from the same country (relative to the total number of references to the work of universities, regardless of location).

The literature about proximity tells us two things. First, two firms at the same distance from a university will have the same probability to cite the university. Hence, Hypothesis 1 means that for instance University of Bourgogne (France) may be cited by companies in Paris (France) as much as by their counterparts in the Luxembourg City (Luxembourg), because both Paris and Luxembourg City are around 300 km away from the University of Bourgogne. Despite being at the same distance, the citation is national in the French case and international in the Luxembourgish case. This is likely to occur for many other universities which are at the same distance from Paris and Luxembourg City because France is larger than Luxembourg.

Second, the literature about proximity tells us that citations occur with a geographical decay. Hence, firms in Paris and Luxembourg City will tend to cite the University of Bourgogne less than other universities which are not so far away, but firms in Paris are likely to find such universities in France whereas Luxembourgish firms are not so likely to find them in Luxembourg, again because France is larger than Luxembourg.

For both reasons, we expect the share of citations to national universities to be higher in France than in Luxembourg. Notice that the literature about proximity usually considers that size is geographic, but we will also deal with demographic, economic and scientific size. Guellec and van Pottelsberghe (2001) provided related evidence showing that smaller countries are more internationalised technologically.

Localisation of knowledge flows and national research structure

Hypothesis 1 is based on visualizing the size of a country in absolute terms. We next apply some relative reasoning, based on the institutional structure of a country's research. If there is geographic localisation of knowledge flows based on patent evidence, this implies that it is evidence of firms' scientific activities because patents are mostly assigned to firms (institutions such as universities and public research organisations, have less of a tradition of using this mechanism of protection). Hence, it follows that the stronger the research activity of the domestic firms relative to other institutions, the more localised will be the country's knowledge.

Recall that most of the existing evidence is based mainly on contexts where firms perform a large share of R&D, mainly the US; only one case refers to European countries (Maurseth and Verspagen 2002). However, when we look at a country's aggregate patent citations what we are observing is patent citations from the regions where firms perform a large share of R&D and patents will include a higher number of citations (Acosta and Coronado 2003). So the claim that knowledge is localised should be interpreted as meaning that knowledge is localised in contexts where firms perform a large share of R&D and there is evidence that in contexts knowledge where firms perform a small share of R&D is delocalised: Azagra-Caro et al. (2009) explain this as follows. In regions where firms perform a small share of R&D, local patents have a low technological profile, so citations to the state-of-the-art will be found elsewhere and firms will find it more difficult to justify a certain degree of novelty and develop more international search strategies. Hence, the predominance of business R&D would appear to be a condition for the localisation of knowledge, which leads to Hypothesis 2:

Hypothesis 2. The higher the share of a country's research that is performed by firms, the higher the share of references to the work of universities from the same country (relative to the total number of references to the work of universities, regardless of location).

Guellec and van Pottelsberghe (2001), for example, find a negative relation between low technological intensity and internationalisation. However, Hypothesis 2 relates more to the institutional structure of Research and Development (R&D) than its intensity. Abramo et al. (2010), in a study of Italian regions, find that the share of intraregional in total university-industry cooperation is higher in regions where firm and private R&D expenditure is concentrated, although they do not test this quantitatively.

Data and dependent variable

The data collection for the present study was designed by the Institute for Prospective Technological Studies (IPTS) in 2009. An international consortium of researchers from the University of Newcastle, Incentim and the Centre for Science and Technology Studies (CWTS) were responsible for implementing the data collection.

The EPO Worldwide Patent Statistical Database (PATSTAT) database was used to compile a dataset of 649,156 direct EPO patents applied for in the period 1990-2007. These were classified by applicant country. In the case of multiple countries, fractional counts were applied, i.e. if a patent application involved two different countries, each scored 0.5 patents.

These 649,156 patents involved 1,938,818 references, equating to an average of 3 references per patent (cf Criscuolo and Verspagen 2008 and Sapsalis et al. 2007). We then identified which were university references. The strategy used differed depending on whether it was references to patent literature or to non-patent literature.

For patent literature, the procedure adopted was identifying references to patents with at least one university listed as an applicant, i.e. university-owned patents. What the literature refers to as university-invented patents (patents applied for by non-university actors, with university inventors) are not part of this study. This would be a problem if the aim were to measure references to all university patenting. However, our target is access to university public knowledge bases, which is better represented by university-owned patents, because most university-invented patents belong to the private sector. The data source was PATSTAT and data gathering was exhaustive.

For the non-patent literature, the adoption of certain criteria led to some exclusions. First, we include only scientific references to documents included in the Web of Science (WoS) database, especially research-based documents, so-called ‘research articles’, ‘research reviews’, ‘letters’ and ‘notes’. This has some limitations in terms of coverage of scientific fields, English language bias, etc., but the quality of the data is widely acknowledged. Second, the sample is restricted to single authored documents, which may raise more serious concerns. On the one hand, it implies a major underestimation of university-technology links since many university papers are co-authored. Our response is that the aim of this paper does not depend on the precise value of these links, but on the calculation of an average to analyse evolution and variation by country. On the other hand, the single-author criterion could introduce bias if, in some years or in some countries there is a disproportionately higher share of co-authored papers in total university papers. Our response to this is that we assume that the distributions of single-authored and co-authored university papers across countries and years are similar, which is reasonable, because if authors from nationalistic and non-nationalistic countries are randomly distributed, the positive and negative errors in the estimated regressions will cancel each other out. In addition, we will include country and year fixed effects to control for unobserved heterogeneity. Nevertheless, we provide a breakdown of the data by patent technology class to check the robustness of our results.

These matching procedures for the distribution of references by institutional sector resulted in 82% non-university references, 17% references of unknown institutional origin and 1% university references. As explained above, this 1% is an underestimation due to the single-author criterion.

This 1%, or 20,630 university references (contained in 15,433 patents), is the basis for our analysis. Among these, the distribution by type of literature is: 67% patent literature (i.e. university-owned patents) and 33% non-patent literature (i.e. university-authored papers). The latter percentage is again an underestimation due to the single-author criterion.

University references are classified by country of the university. Based on our classification by country for patent applicants we are able to check whether there is a match between applicant country and country of cited university. The resulting figures are 90% international university references and 10% national university references (0.2% could not be assigned).

This 10% is the average value of the main variable of interest in this paper:

- SNAT: share of national university references over total number of university references

Descriptive results

The following description applies to the 37 countries in Eurostat's R&D statistics, which account for 99.9% of the patents in the database.

The 10% average of SNAT is fairly stable in the 18 year period studied: 1990-2007. The trend depicted in Fig. 1 may appear to oscillate, but is roughly maintained at 10% with no sustained peaks or declines. Hence, time variation does not appear to be important.

{Fig. 1 around here}

However, country variation is substantial. Fig. 2 shows that non-EU countries account for a larger share of national university references than EU countries. This is due to the very large shares of the US and the Russian Federation, both of which are large countries; this result is in line with our theory. Within the EU, the UK's first ranking also satisfies the theory, and is based on its high level of business R&D. The Scandinavian countries are below the average, which is fairly predictable because they are small countries although they have several science-intensive firms. Most of the more recent EU member states are ranked rather low, which again is as expected since all these countries are relatively small and have fewer resources for firm R&D.[1]

{Fig. 2 around here}

These results suggest that in order to test Hypothesis 1 and Hypothesis 2 in an econometric setting, it is convenient to control for country heterogeneity. The analysis also controls for time variation, although this is likely to be not significant.

[1] Notice that the countries included in the figures are those with R&D statistics in Eurostat, i.e. the 27 EU member states plus ten non-EU countries. Countries like China and Norway appear in the graph because they report R&D statistics to Eurostat, even if SNAT=0. On the contrary, other countries are not there either because they do not have EPO patents in PATSTAT (e.g. Pakistan, Bangladesh, etc.) or do have them but their R&D statistics do not appear in Eurostat (e.g. India, Indonesia, etc.)

Econometric methods

If we multiply the years (18) by the 37 countries in Eurostat's R&D statistics we have an initial panel of 666 observations, where the unit of analysis is country-year.

We define this variable as:

- UNIVREF: number of university references.

Note that since SNAT is a ratio whose denominator is UNIVREF, SNAT exists only if UNIVREF is positive. This applies to 55% of cases, which means the panel for our estimations drops to 369 observations. This may introduce a sample selection problem, i.e. it may exclude some countries and years, which may bias results. We address this problem later in the paper (it is found not to be an issue). In the meantime, the model to be estimated is:

$$SNAT_{it} = f\,(size_{i,t-1}, researchstructure_{i,t-1}) \qquad (1)$$

The subscripts i and t respectively stand for country and year. Years are lagged one period on the right-hand side to reduce endogeneity.

The variable SNAT is continuous and presents a large proportion of zero values (47%), which may be the outcome of two different distributions: first, the decision of whether or not to insert a national university reference; and second, the decision related to the actual number of national university references to insert. Hence, SNAT may be censored and Tobit models of estimation would seem adequate.

The size variables allow us to test Hypothesis 1, which predicts a positive sign of the coefficients. The conception of 'size' in this study encompasses different dimensions of size: geographic, demographic, economic and scientific size:

- Surface area (SURFACE), from the United Nations' Demographic and Social Statistics;

• Population (POP), from the United Nations' World Population Prospects;

• Gross domestic product (GDP), from Eurostat;

• Gross domestic expenditure on R&D (GERD) from Eurostat's R&D statistics[2].

The research structure variables allow us to test Hypothesis 2. They are:

• SBERD: share of Business expenditure on R&D (BERD) over GERD;

• SHERD: share of Higher education expenditure on R&D (HERD) over GERD;

• SGOVERD: share of Government expenditure on R&D (GOVERD) over GERD;

• SPNPERD: share of Private non-profit institutions expenditure on R&D (PNPERD) over GERD.

According to Hypothesis 2, we can expect a positive sign for the coefficient of share of business R&D.

Data on R&D and GDP have many missing values. In order to add some information, for missing years, we used the average value of the year before and the year after where available. This generated another 5% of records. We also had the problem that downloads from Eurostat generate a ":" sign for both missing and zero values. Wherever possible and applicable, ":" was replaced by 0. In practice, this was an issue only with PNPERD, for which a high number of recoveries was possible. Nevertheless, the panel reduced from 369 to 315-323 observations (however, later models show that the results are consistent with larger numbers of observations).

Table 1 shows the sample for the panel, i.e. observations where SNAT can be calculated (with a positive number of university references or UNIVREF=0). Average SNAT is the 10% identified in the descriptive results. Surface is in millions of square

[2] We also tried each one of the four components of GERD by institutional sector: Business expenditure on R&D (BERD), Higher education expenditure on R&D (HERD), Government expenditure on R&D (GOVERD) and Private non-profit institutions expenditure on R&D (PNPERD). All produced the same results (available on request).

kilometres. Population is in thousands of millions inhabitants. GDP and R&D variables are in billions of euro. The sum of the structural variables (SBERD+SHERD+SGOVERD+SPNPERD) logically equals 1.

{Table 1 around here}

Table 2 shows that the size variables are all highly correlated, with correlation coefficients close to 1. From an econometric point of view, it is not correct to estimate them in the same equation because one simply absorbs the effect of the other and the significance and signs change artificially. The approach followed is to estimate separate regressions for each and refine the model with a better fit. SBERD is also highly correlated with SHERD and SGOVERD, so the latter are used as the benchmark.

{Table 2 around here}

Hence, the initial models to be estimated are:

$$SNAT_{it} = f\,(SURFACE_{i}, SBERD_{i,t-1}, SPNPERD_{i,t-1}) \tag{2}$$

$$SNAT_{it} = f\,(POP_{i,t-1}, SBERD_{i,t-1}, SPNPERD_{i,t-1}) \tag{3}$$

$$SNAT_{it} = f\,(GDP_{i,t-1}, SBERD_{i,t-1}, SPNPERD_{i,t-1}) \tag{4}$$

$$SNAT_{it} = f\,(GERD_{i,t-1}, SBERD_{i,t-1}, SPNPERD_{i,t-1}) \tag{5}$$

Country and year fixed effects are added to the best-fit model. However, including all country fixed effects produces multicollinearity. We solved this by grouping countries in meaningful blocks in order to find a parsimonious model with more degrees of freedom. The country groupings are:

- US
- EU06: Belgium, France, Germany, Italy, Luxembourg, Netherlands

- EU12: United Kingdom, Denmark, Ireland, Greece, Portugal, Spain
- EU15: Austria, Finland, Sweden
- EU27: Cyprus, Czech, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia, Bulgaria, Romania
- Other non-EU countries: China, Croatia, Iceland, Japan, Norway, Russia, South Korea, Switzerland, Turkey.

EU12 is the benchmark category for the estimations. Other combinations of blocks (e.g. separating EU09 from EU12, or EU25 from EU27) led to the same results – these are presented in the next section.

Econometric results

Aggregate evidence showing that the higher the share of business R&D, the higher is the share of national university references

Table 3 shows the results of the Tobit estimation of SNAT as a dependent variable. Columns 1 to 4 of Table 3 present the regressions using different size variables (equations 2 to 5). The four size variables are positive and significant, which is evidence supporting Hypothesis 1. All regressions but the third have the same number of observations, so the DECOMP based fit measure can be seen as confirming that the regression with SURFACE as the size variable is the best fit.[3] The third regression uses GDP as the size measure and has a different number of observations, so the DECOMP based fit measure does not allow for comparison. The Bayesian Information Criterion

[3] The DECOMP based fit measure is a pseudo R^2, calculated as the variation of the predicted mean relative to the observed mean divided by the sum of the numerator and a residual variance of the true value minus the conditional mean function (Greene 2002).

(BIC) is more useful. BIC is lower in regression 1 than in regression 3, so SURFACE is preferable to GDP as a size measure. SURFACE, therefore, is the best predictor of SNAT.

{Table 3 around here}

In the first four regressions, the two research structure variables (SBERD and SPNPERD) show consistent results: SBERD is always positive and significant, which provides support for Hypothesis 2; SPNPERD is never significant.

Regression 5 includes country block and time fixed effects. Their inclusion makes the coefficient of SURFACE insignificant, which does not support Hypothesis 1, although Hypothesis 2 still holds. The time effects are also not significant, which is consistent with Fig. 1 in the section on Descriptive results; likelihood ratio tests reveal that time effects can be removed and that the best model is the model including only country block effects.

This is used for regression 6. Regression 6 shows the best fit in terms of the BIC measure. It confirms that there is a lack of evidence in support of Hypothesis 1 but that Hypothesis 2 holds.

We had reasons to suspect that Hypothesis 1 would not hold: language bias (Meyer 2000), national differences in exploration patterns (Rosenkopf and Nerkar 2001), degree of social proximity (Breschi and Lissoni 2005; Agrawal et al. 2006; 2008), interpersonal ties (Singh 2005), breakthrough innovations (Phene et al. 2006), etc are sources of delocalisation[4]. The country block fixed effects may be a better proxy for these phenomena. A closer look at the country dummies is interesting: US and older EU

[4] This points to a lesser, but not necessarily non-existent influence of localisation on knowledge flows, which leaves the debate open. Some studies hypothesise about the increasing importance of geography in knowledge flows measured by through patent citations (Sonn and Storper 2008).

member states score similar in SNAT, whereas the newest member states include a higher share of international university references; non-EU countries other than the US, have low coefficients. This suggests that integration processes (with the US - a political union of many states - at the forefront) counterbalance the influence of the sources of delocalisation.

Robustness check 1: controlling for sample selection

In the Methodology and data section, we showed that the models in Table 3 contain only those observations with a positive number of university references, i.e. where UNIVREF>0. This may create a sample selection problem due to the omission of relevant observations, which could bias the results. One way to resolve this problem is to create the following variable:

- UNIVREFP: equal to 1 if UNIVREF is positive and 0 otherwise;

and to run a regression on its determinants, which we use as a selection equation for the Tobit estimation.

We created UNIVREFP (average equal to 0.55, i.e. 55% of the observations have a positive number of university references) and ran some regressions using the same determinant as for SNAT. The results are included in the appendix. Using the best fit model for UNIVREFP as a selection equation, we re-estimated the determinants of the share of national university references.

The results are shown in Table 4, regression 1. Although very similar to the benchmark (regression 6 in Table 3), note that there is no sample selection (because the ρ parameter is not significant). Hence, it is not worth pursuing sample selection further; on the contrary, we can confirm that the benchmark is a better model. However, we can perform another robustness check.

{Table 4 around here}

Note that in the sample selection model, we do not use the original variable SNAT but include a transformation SNAT2, which is equal to 0 if UNIVREF=0 and to SNAT in every other case. Logically, the number of observations is larger (increasing from 323 to 527). It is worth investigating whether the determinants of SNAT2 are the same as for SNAT.

Table 4, regression 2 provides this information. We repeated all the processes described in Table 3: we selected the best size predictor –SURFACE– and included fixed effects –time effects were not significant. In the final model, there is no evidence to support Hypothesis 1 but there is support for Hypothesis 2. So even using the transformed dependent variable the results are the same.

Robustness check 2: breaking down the sample by technology classes (IPC sections)

In the section on Data and the dependent variable we referred to an important limitation of the database which is its restriction to single authored university papers. We tackle this problem indirectly by showing the data broken down by technology class. Our aim is to show that the results are consistent: if excluding multi-authored university papers imposes a severe bias, it would show up no matter how the data were broken down.

For technology class, we use the World Intellectual Property Organization's (WIPO) eight International Patent Classification (IPC) sections –see list in Table 5. Of course, the number of patents with university references per class decreases as does the number of observations of the dependent variable SNAT: the maximums are for A. Human Necessities, C. Chemistry, Metallurgy and G. Physics; the minimum is for D. Textiles, Paper and E. Fixed Constructions.

{Table 5 around here}

In order to have a reasonable number of observations (i.e. more than 200), we include the eight IPC sections in three groups, following Azagra et al. (2006), which roughly links IPC sections to Pavitt's (1984) sectoral taxonomy. While sections A, D and E are more likely to apply to supplier dominated sectors, sections B and F are more likely to apply to production intensive sectors, and sections C, G and H to science-based sectors.

For each of the three groups, we apply the same estimation protocol as before: SURFACE is always the best predictor of size, and time effects can always be dropped. This produces the full models depicted in Table 6, which we reduce by removing time effects.

{Table 6 around here}

We can see that, again, Hypothesis 2 holds for each group of technologies, and that higher shares of BERD are associated with higher shares of national university references.

Hypothesis 1 is confirmed in one of the three cases: for IPC sections related to production intensive sectors (B and F). Hence, in these technologies, which is against the aggregate trend, larger countries present higher shares of national university references. This could reveal a weakness in the data related to the exclusion of multiple authored university papers. However, there is a reason to think that the result is reasonable: tacit knowledge is more important in industries such as machinery or automobiles, than in traditional or high-tech industries (where explicit knowledge is more important). This means that the former types of industry are characterised by a more local knowledge frontier than the latter (Breschi and Malerba 1997). Hence, the finding that country size has more influence on citations to national university references for IPC sections related to production intensive sectors than for other technology classes, can be accepted as reasonable.

Limitations and suggestions for future research

The discussion on knowledge sharing narrowed to a focus on one facet: access to university knowledge. Therefore, the results cannot be extended to the wider phenomenon of knowledge sharing because they are not related to other aspects, such as harmonised intellectual property rights, shared principles for knowledge transfer, cooperation between public research and industry, communication to the public of scientific knowledge, etc., which would require a different theoretical basis and different indicators. However, for university-industry cooperation, research shows a positive association between higher levels of regional business R&D and participation in regional joint projects (Azagra et al. 2011).

Access to university knowledge bases can be measured in other ways than by the measure of knowledge flows applied in this paper; for instance, university references in scientific papers, especially firm authored papers (see e.g. Tijssen and Van Leeuwen 2006). Applying this measure would provide interesting comparisons. Even within our own measure, SNAT, it is possible to disaggregate references to university-owned patents and university (single authored) papers. Distinguishing between both might result in additional insights, but it would require its own theoretical development (i.e. making hypotheses about whether the expected impact of the independent variables will be different on each type of reference).

The major limitation of the empirical approach in this paper is the focus only on single-authored university papers. To what extent their country distribution is representative of the whole population could be questioned, although so far, there is no evidence to challenge this claim, and no substantive reason for it to invalidate our theory. To test the robustness of our results, we tried different breakdowns of the data and estimation techniques. A more complete analysis would include all university papers in the model, and this could be the subject of further research.

Comparison with non-university references found in patent documents would be another natural extension of this work. Do they follow the same geographical patterns as the ones identified in this paper? Constructing these data is beyond the scope of this research.

Conclusions

The first –theoretical– contribution of this paper is to show that access to the public knowledge base varies across countries as a positive function of country size and the relative share of business funding of R&D. We used countries as the unit of observation in the analysis of knowledge flows through patent citations which allows for a more contextualised interpretation of the results for localisation of knowledge flows.

The second contribution of this paper is an institutional approach to patent citation analysis to identify university references. The existing work distinguishes between patent literature (to measure knowledge spillovers) and non-patent literature (to measure science-technology interactions). The approach in this paper is more appropriate to address current policy concerns related to public access to the knowledge produced by universities.

The third contribution of this paper is the construction of indicators for access to university knowledge. We provide three main findings.

First, the indicators show that there is international access to the university knowledge base and it reached a plateau at least from the 1990s. Other indicators produce similar results (Ponds 2009). An implication for policy is the need to understand the limits of integration. Those keen to make the idea of an ERA a reality by coordinating national research policies should note that the increased coordination in the last two decades has had no impact on changing the current plateau. It is therefore questionable whether further efforts at coordination would yield significant results in terms of integration related to of access to university knowledge. More coordination might be justified in terms of more integrated research markets (more opportunities for researchers), dissemination of science to society, etc.

A second empirical finding is the high level of variation across countries in the composition of access to the university knowledge base, within and across borders. Policy should try to refine objectives: given that current policy discourse favours both local and global contributions of universities, more precise targets are needed. The question should perhaps be not how to increase access, but rather what should be the research focus of individual countries.

The third empirical finding is that increases in the national share of business funding of R&D enhance access to domestic university knowledge. The policy implication is that nationalist countries should focus on increasing BERD rather than other components of GERD, which probably means a reversal of current trends: since 1990, growth of HERD and GOVERD has been higher on average than growth of BERD. For internationalist countries, the current trends should be sustained. Perhaps being nationalist or internationalist in more convenient depending on the technology life cycle, since this affects the impact of local interaction on technological performance (Lecocq and Van Looy 2009). We hope that the findings in this paper provide some hints about achieving the various goals referred to.

Acknowledgements This research was initiated with the framework of ERAWATCH, a joint initiative of the European Commission's Directorate General for Research and the Joint Research Centre-Institute for Prospective Technological Studies (IPTS). The views expressed in this article are those of the author and do not necessarily reflect those of the European Commission (EC). Neither the EC nor anyone acting on behalf of the EC is responsible for the use that might be made of the information. I am grateful to René van Bavel and Xabier Goenaga for their support and to Laura de Dominicis for exchange of ideas. I am also grateful to the international consortium that produced the database, including Henry Etzkowitz, Marina Ranga and members of Incentim and CWTS, led, respectively, by Bart Van Looy and Robert J.W. Tijssen. Previous versions of the paper were presented at the Triple Helix VIII International Conference on University, Industry and Government Linkages and the IPTS Workshop "The Output of R&D activities: Harnessing the Power of Patents Data – II", and I acknowledge helpful comments from the audiences. My colleagues at INGENIO also provided useful comments on a seminar presentation.

Appendix: estimating the determinants of a selection equation with UNIVREFP as a dependent variable

Table 7 presents the results, following the logic applied in Table 3. The χ^2 test indicates that all the models are significant.

{Table 7 around here}

Each of the first four regressions in Table 7 includes one independent size variable. The fit, according to the percentage of correct predictions, is practically the same in all of them; but the BIC is lower (better) for regression 3 which uses GDP, so we use this as the basis for the remaining ones. SBERD and SPNPERD are always significant.

Regression 5, using GDP, includes country block and time effects. Only one time effect, for year 2007, is significant (not shown). Hence, in regression 6, we include only country block effects. However, the likelihood ratio test indicates that regression 5 is preferred –time effects are necessary. In an attempt to reduce multicollinearity and gain degrees of freedom, we use another strategy: to find a reduced model with significant variables only. To do this, we keep rerunning the model, each time dropping the least significant variable, until we are left with only significant variables. This process leads to regression 7, with fixed effects for year 2007 only (a drop in the number of university references due to a similar drop of patents in the original data set, because of the lag in the introduction of data). The sign and significance of the remaining coefficients is consistent in all the models, including regression 7. Therefore, we use regression 7 as the selection equation for the Tobit model.

List of acronyms

BERD: business expenditure on R&D

BIC: Bayesian Information Criterion

CWTS: Centre for Science and Technology Studies

EPO European Patent Office

ERA: European Research Area

EU: European Union

GDP: gross domestic product

GERD: gross domestic expenditure on R&D

GOVERD: government expenditure on R&D

HERD: higher education expenditure on R&D

IPC: International Patent Classification

IPTS: Institute for Prospective Technological Studies

PATSTAT: EPO Worldwide Patent Statistical Database

PNPERD: private non-profit institutions expenditure on R&D

POP: population

R&D: research and development

SBERD: share of BERD over GERD

SHERD: share of HERD over GERD

SGOVERD: share of GOVERD over GERD

SNAT: share of national university references over total number of university references

SNAT2: a variable equal to 0 if UNIVREF=0 and to SNAT in every other case

SPNPERD: share of PNPERD over GERD

UK: United Kingdom

UNIVREF: number of university references

UNIVREFP: a variable equal to 1 if UNIVREF is positive and 0 otherwise

US: United States

USPTO: US Patent and Trademark Office

WIPO: World Intellectual Property Organization

WoS: Web of Science

References

Abramo, G., D'Angelo, C.A., & Solazzi, M. (2010). Assessing public–private research collaboration: is it possible to compare university performance? *Scientometrics, 84*, 173–197.

Acosta, M. & Coronado, D. (2003). Science-technology flows in Spanish regions: An analysis of scientific citations in patents. *Research Policy*, *32(10)*, 1783–1803.

Acosta, M., Coronado, D., Ferrándiz, E., & León, M.D. (2011). Factors affecting inter-regional academic scientific collaboration within Europe: the role of economic distance. *Scientometrics, 87*, 63–74.

Agrawal, A., Cockburn, I. & McHale, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography, 6(5)*, 571–591.

Agrawal, A., Kapur, D. & McHale, J. (2008). How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics, 64(2),* 258-269.

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

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

Azagra-Caro, J.M., Fernández-de-Lucio, I., Perruchas, F. & Mattsson, P. (2009). What do patent examiner inserted citations indicate for a region with low absorptive capacity? *Scientometrics, 80(2),* 441-455.

Azagra-Caro, J.M., Pontikakis, D. & Varga, A. (2011). Delocalisation patterns in University-Industry interaction: Evidence from the 6th R&D Framework Programme, *European Planning Studies*, forthcoming.

Azagra-Caro, J.M., Yegros-Yegros, A. & Archontakis, F. (2006). What do university patent routes indicate at regional level? *Scientometrics, 66(1),* 219-230.

Breschi, S. & Lissoni, F. (2005). Knowledge networks from patent data. *Handbook of quantitative science and technology research*, 613-643.

Breschi, S. & Lissoni, F. (2001). Knowledge spillovers and local innovation systems: a critical survey. *Industrial and corporate change, 10(4),* 975-1005.

Breschi, S. & Malerba, F. (1997). Sectoral Innovation Systems: Technological Regimes, Shumpeterian Dynamics, and Spatial Boundaries, in C. Edquist (ed.): *Systems of Innovation: Technologies, Institutions and Organizations*, ch. 6. Londres and Washington: Pinter

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

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

EC (2007). *Commission Green Paper 'The European Research Area: New Perspectives'*, COM(2007) 161.

Feldman, M.P. (1999). The New Economics of Innovation, Spillovers and Agglomeration: Ar eview of Empirical Studies. *Economics of Innovation and New Technology, 8(1),* 5-25.

Greene, W.H. (2002*). LIMDEP Version 8.0 econometric modeling guide 2*, Econometric Software, Inc., Plainview, NY.

Guellec, D & van Pottelsberghe de la Potterie, B. (2001). The internationalisation of technology analysed with patent data. *Research Policy, 30,* 1253–1266.

Hu, A.G.Z. & Jaffe, A.B. (2003). Patent citations and international knowledge flow: the cases of Korea and Taiwan. *International Journal of Industrial Organization, 21(6)*, 849-880.

Jaffe, A.B. & Trajtenberg, M. (1999). International knowledge flows: evidence from patent citations. *Economics of Innovation and New Technology, 8(1),* 105-136.

Jaffe, A.B. & Trajtenberg, M. (1996). Flows of knowledge from universities and federal labs: modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences, 93,* 12671–12677.

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

Lecocq, C., & Van Looy, B. (2009). The impact of collaboration on the technological performance of regions: time invariant or driven by life cycle dynamics? *Scientometrics, 80(3),* 847–867.

Maurseth, P.B. & Verspagen, B. (2002). Knowledge spillovers in Europe: a patent citations analysis. *Scandinavian Journal of Economics, 104(4),* 531-545.

Meyer, M. (2000). Does science push technology? Patents citing scientific literature. *Research Policy, 29(3)*, 409-434.

Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy, 13,* 343-373.

Phene, A., Fladmoe-Lindquist, K. & Marsh, L. (2006). Breakthrough innovations in the US biotechnology industry: the effects of technological space and geographic origin. *Strategic Management Journal*, *27(4),* 369-388.

Ponds, R. (2009). The limits to internationalization of scientific research collaboration. *Journal of Technology Transfer, 34*, 76-94.

Rosenkopf, L. & Nerkar, A. (2001). Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal, 22(4),* 287-306.

Sapsalis, E. & van Pottelsberghe de la Potterie, B. (2007). The institutional sources of knowledge and the value of academic patents. *Economics of Innovation and New Technology, 16(2),* 139-157.

Singh, J. (2005). Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science, 51(5),* 756-770.

Sonn, J.W. & Storper, M. (2008). The increasing importance of geographical proximity in technological innovation: an analysis of US patent citations, 1975-1997. *Environment and Planning, 40(5),* 1020-1039.

Tijssen, R.J.W. (2001). Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. *Research Policy, 30(1),* 35-54.

Tijssen, R.J.W. & Van Leeuwen, T.N. (2006). Measuring impacts of academic science on industrial research: A citation-based approach, *Scientometrics, 66(1),* 55-69.

Verbeek, A., Debackere, K., & Luwel, M. (2003). Science cited in patents: A geographic “flow” analysis of bibliographic citation patterns in patents. *Scientometrics, 58(2),* 241-263.

Figures

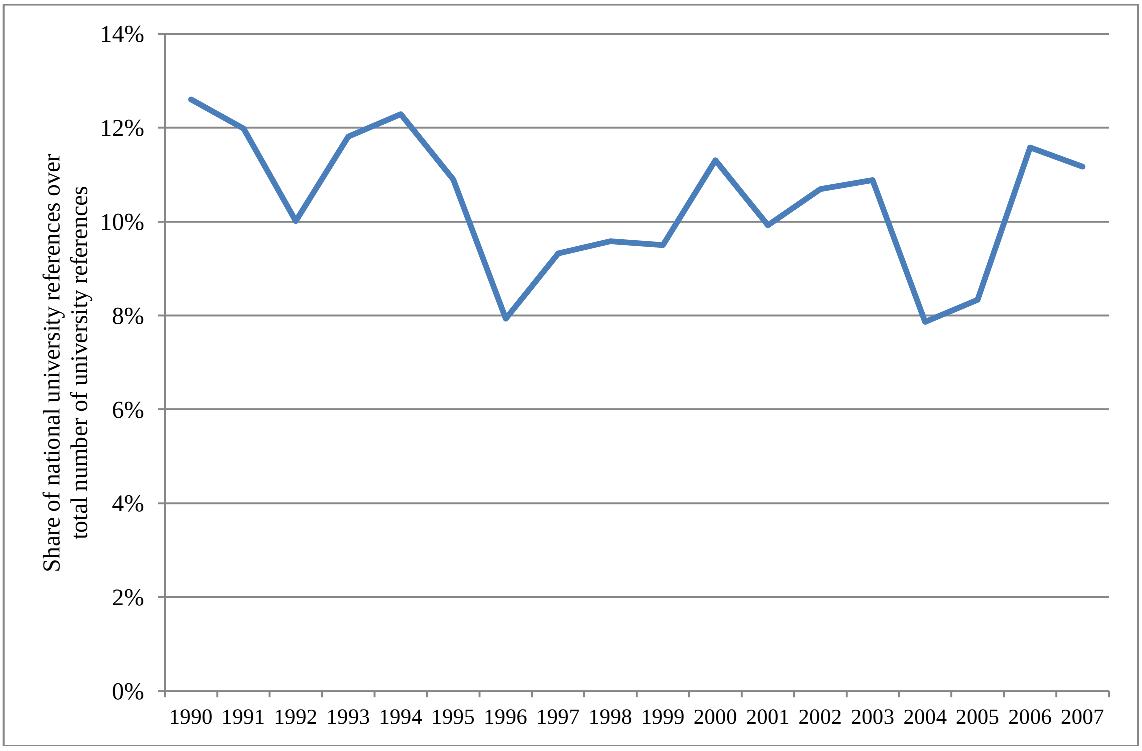

Fig. 1 The time stability in the share of national university references around 10%

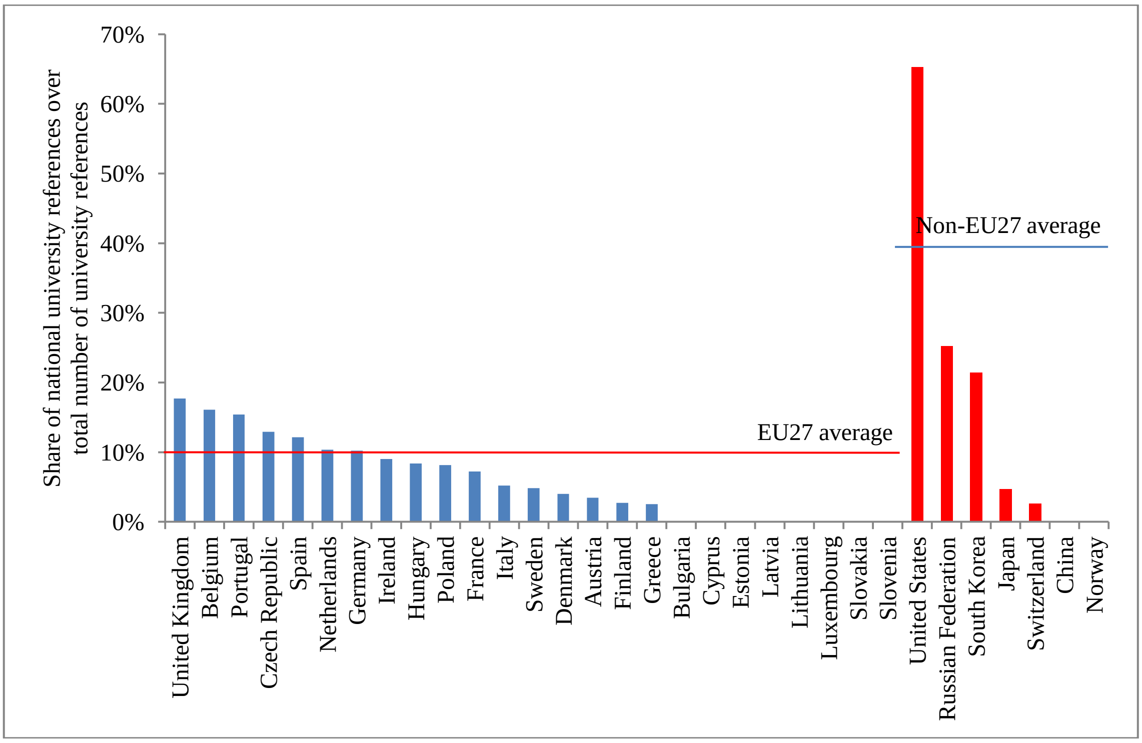

Fig. 2 The large variation across countries in the share of national university references. Countries are ordered first by block (EU, non-EU), then in decreasing order of SNAT.

Tables

Table 1 Descriptive statistics (sample with UNIVREF>0)

	Mean	Standard deviation	Minimum	Maximum	Number of cases
SNAT	0.10	0.18	0.00	1.00	369
SURFACE	0.85	2.66	0.00	17.10	369
POP	0.04	0.08	0.00	1.15	369
GDP	1.01	1.87	0.01	11.46	325
GERD	0.02	0.05	0.00	0.31	333
SBERD	0.59	0.15	0.19	0.93	323
SHERD	0.23	0.08	0.00	0.51	323
SGOVERD	0.16	0.09	0.01	0.71	323
SPNPERD	0.02	0.04	0.00	0.28	323

Table 2 Correlation matrix

	SURFACE	POP	GDP	GERD	SBERD	SHERD	SGOVERD	SPNPERD
SURFACE	1.00							
POP	0.88	1.00						
GDP	0.85	0.96	1.00					
GERD	0.83	0.94	0.99	1.00				
SBERD	0.20	0.26	0.29	0.31	1.00			
SHERD	-0.31	-0.39	-0.38	-0.40	-0.74	1.00		
SGOVERD	-0.08	-0.11	-0.17	-0.18	-0.80	0.23	1.00	
SPNPERD	0.12	0.10	0.12	0.14	-0.43	0.15	0.26	1.00

Table 3 Tobit models with SNAT as a dependent variable

	1	2	3	4	5	6
Number of observations	323	323	315	323	323	323
Log likelihood function	-25	-31	-44	-56	8	2
DECOMP based fit measure	0.53	0.44	0.39	0.36	0.50	0.50
	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)
Constant	-0.14 (-2.59) ***	-0.18 (-3.01) ***	-0.15 (-2.24) **	-0.14 (-2.02) **	-0.12 (-1.5)	-0.13 (-1.93) *
SURFACE	0.07 (14.7) ***				0.08 (1.44)	0.08 (1.39)
POP		2.33 (13.31) ***				
GDP			0.07 (11.13) ***			
GERD				2.55 (10.22) ***		
SBERD	0.23 (2.74) ***	0.19 (2.07) **	0.19 (1.82) *	0.19 (1.76) *	0.25 (2.52) **	0.25 (2.46) **
SPNPERD	0.02 (0.07)	0.02 (0.05)	-0.05 (-0.09)	-0.15 (-0.36)	0.45 (1.22)	0.5 (1.34)
Country block effects	Not included	Not included	Not included	Not included	Included	Included
US					-0.15 (-0.28)	-0.13 (-0.25)
EU06					0.04 (1.47)	0.04 (1.41)
EU15					-0.13 (-3.56) ***	-0.12 (-3.43) ***
EU27					-0.17 (-3.99) ***	-0.16 (-3.9) ***
Other non-EU countries					-0.13 (-2.95) ***	-0.13 (-2.9) ***
Time effects	Not included	Not included	Not included	Not included	Included	Not included
σ	0.17 (17.94) ***	0.18 (18.23) ***	0.19 (18.01) ***	0.2 (18.07) ***	0.16 (18) ***	0.16 (18) ***
BIC	66.36	79.77	105.24	129.09	127.57	42.20

***1% significant. ** 5% significant. * 10% significant.

Table 4 Tobit models with SNAT2 as a dependent variable

	1 Tobit with sample selection	2 Tobit without sample selection
Number of observations	527	527
Log likelihood function	-158	-39
DECOMP based fit measure	-	0.45
	Coeff. (t-ratio)	Coeff. (t-ratio)
Constant	-0.15 (-1.28)	-0.23 (-3.94) ***
SURFACE	0.09 (0.98)	-0.01 (-1.11)
SBERD	0.28 (1.75) *	0.41 (4.45) ***
SPNPERD	0.78 (1.58)	0.69 (1.79) *
US	-0.22 (-0.26)	0.7 (5.16) ***
EU06	0.04 (0.96)	0.06 (2.12) **
EU15	-0.12 (-2.79) ***	-0.11 (-2.94) ***
EU27	-0.14 (-2.02) **	-0.27 (-6.92) ***
Other non-EU countries	-0.13 (-1.75) *	-0.24 (-5.77) ***
σ	0.16 (23.52) ***	0.18 (17.75) ***
ρ	-0.06 (-0.13)	

***1% significant. ** 5% significant. * 10% significant.

Table 5 Breakdown of patents with university references by IPC section

IPC section	IPC name	Number of patents with at least one university reference
A	Human Necessities	3,817
B	Performing Operations; Transporting	1,250
C	Chemistry; Metallurgy	3,831
D	Textiles; Paper	107
E	Fixed Constructions	140
F	Mechanical Engineering; Lighting; Heating; Weapons; Blasting	502
G	Physics	3,447
H	Electricity	2,333
	Not assigned	7
	Total	15,433

Table 6 Tobit models with SNAT as a dependent variable, by IPC groups of sections

	1 IPC sections related to supplier dominated sectors (A, D and E)		2 IPC sections related to production intensive sectors (B and F)		3 IPC sections related to science-based sectors (C, G and H)	
	Full model	Reduced model	Full model	Reduced model	Full model	Reduced model
Number of observations	274	274	212	212	301	301
Log likelihood function	11	1	-67	-71	-40	-49
DECOMP based fit measure	0.47	0.46	0.44	0.44	0.40	0.40
	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)
Constant	-0.19 (-1.94) *	-0.19 (-2.18) **	-0.53 (-2.23) **	-0.64 (-2.82) ***	-0.13 (-1.21)	-0.12 (-1.28)
SURFACE	0.03 (0.55)	0.04 (0.61)	0.28 (2.2) **	0.27 (2.15) **	0.1 (1.27)	0.09 (1.19)
SBERD	0.31 (2.39) **	0.31 (2.32) **	0.87 (2.71) ***	0.9 (2.78) ***	0.29 (2.02) **	0.25 (1.75) *
SPNPERD	-0.33 (-0.54)	-0.24 (-0.4)	2.04 (0.99)	2.09 (1.01)	0.6 (1.2)	0.63 (1.24)
US	0.32 (0.58)	0.27 (0.46)	-2.06 (-1.74) *	-2 (-1.67) *	-0.33 (-0.46)	-0.28 (-0.38)
EU06	0.08 (3.1) ***	0.08 (3) ***	-0.03 (-0.58)	-0.02 (-0.39)	0.02 (0.63)	0.02 (0.54)
EU15	-0.1 (-2.74) ***	-0.1 (-2.55) **	-0.48 (-4.61) ***	-0.46 (-4.5) ***	-0.23 (-4.31) ***	-0.22 (-4.11) ***
EU27	-0.23 (-2.96) ***	-0.23 (-2.94) ***	-1.59 (-0.05)	-1.63 (-0.05)	-0.18 (-3.16) ***	-0.18 (-3.1) ***
Other non-EU countries	-0.18 (-2.61) ***	-0.19 (-2.66) ***	-0.38 (-2.64) ***	-0.38 (-2.58) ***	-0.2 (-3.21) ***	-0.19 (-3.04) ***
Time effects	Included	Not included	Included	Not included	Included	Not included
σ	0.16 (16.63) ***	0.16 (16.63) ***	0.28 (12.25) ***	0.29 (12.24) ***	0.21 (17.53) ***	0.21 (17.53) ***

***1% significant. ** 5% significant. * 10% significant.

Table 7 Probit models with UNIVREFP as a dependent variable

	1	2	3	4	5	6	7
Number of observations	527	527	506	527	506	506	506
Log likelihood function	-290	-294	-277	-296	-155	-180	-161
Prob[χ^2 > value]	0	0	0	0	0	0	0
Correct predictions	75%	74%	75%	74%	85%	83%	84%
	Coeff.	Coeff.	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)	Coeff. (t-ratio)
Constant	-1.8 (-8.28) ***	-1.71 (-8.01) ***	-1.7 (-7.34) ***	-1.53 (-7.04) ***	-0.14 (-0.23)	-0.41 (-1.34)	-0.37 (-1.18)
SURFACE	-0.09 (-4.12) ***						
POP		-1.34 (-2.66) ***					
GDP			0.13 (2.08) **		0.43 (3.83) ***	0.31 (3.24) ***	0.41 (3.77) ***
GERD				4.66 (2.09) **			
SBERD	3.83 (10.11) ***	3.65 (9.83) ***	3.41 (8.32) ***	3.11 (8.08) ***	2.97 (5.27) ***	2.51 (4.87) ***	2.8 (5.15) ***
SPNPERD	8.87 (4.66) ***	8.48 (4.5) ***	9.12 (4.07) ***	7.48 (3.97) ***	6.4 (2.6) ***	6.03 (2.58) ***	5.66 (2.37) **
Country block effects	Not included	Not included	Not included	Not included	Included	Included	Included
US					-2.83 (-3.14) ***	-2.28 (-2.94) ***	-2.72 (-3.11) ***
EU06					1.32 (2.23) **	0.83 (2.01) **	1.29 (2.18) **
EU15					1.07 (1.66) *	0.6 (1.32)	1.04 (1.63)
EU27					-1.46 (-6.45) ***	-1.28 (-6.46) ***	-1.36 (-6.46) ***
Other non-EU countries					-2.5 (-8.66) ***	-2.12 (-8.59) ***	-2.34 (-8.69) ***
Time effects	Not included	Not included	Not included	Not included	Included	Not included	Selected
Year 2007							-2.09 (-5.21) ***
BIC	598.37	607.52	572.75	611.78	465.50	409.45	408.94

***1% significant. ** 5% significant. * 10% significant.