Ann Reg Sci (2005) 39:495–522
DOI: 10.1007/s00168-004-0225-6

The Annals of
Regional Science

The effects of scientific regional opportunities in science-technology flows: Evidence from scientific literature in firms patent data

Daniel Coronado, Manuel Acosta

Departamento de Economía General, Facultad de CC. Económicas y Empresariales, Universidad de Cádiz, Duque de Nájera, 8, 11002-Cádiz, Spain. (e-mail: {daniel.coronado,manuel.acosta}@uca.es)

Received: September 2003/Accepted: October 2004

Abstract. Relevant scientific literature has demonstrated that in spaces of smaller scale than the national, the availability of scientific knowledge is also relevant for generating spillover effects that benefit the industrial sector. The proliferation of such literature consistently stressing the importance of physical proximity for the two-way flow of knowledge and for the development and fostering of innovation, together with the high degree of self-government of the Spanish regions (which have the competence to develop their own R&D policies), all suggest that the relationships between the scientific community and the industrial sector may be closer and more productive in the regions where the scientific potential is more relevant, in comparison with other regions. The basic objective of this article is to test for the possible differential effects of a favourable scientific environment on science-technology relationships, and more specifically, to determine if the considerable regional resources directed towards scientific research in local universities are being translated into economic results for industry, by way of better utilisation of scientific knowledge to enable companies to generate more and better innovations in processes and products. The methodology that we employ relates the scientific citations in patent documents - as a basic indicator of these science-technology flows- with various indicators of resources and results of academic research that reflect the scientific research environment. With caution, and recognising the limitations inherent in the NPC (non patent citation) methodology, different econometric specifications permit the conclusion to be drawn that companies of those regions with a more favourable scientific environment make greater use of scientific knowledge.

JEL classification: O31, O38, C21, R59

The authors are members of Eco-21 Research Project. Centra Foundation (Junta de Andalucía). We are grateful for the useful comments on an earlier draft of this paper that were provided by three anonymous referees. This research has been granted by the Ministry of Science and Technology (SEC 2001–3030).

1. Introduction

It is well known that many regional governments in Europe enjoy considerable degrees of autonomy and have set about drawing up plans for R&D. In many of these plans, the financing of scientific work in regional universities constitutes the fundamental part of the planned actions. Confidence that the strengthening of the scientific environment is transferred to industry, by way of more and better innovations of products and processes, has been one of the key arguments in support of allocating huge quantities of resources in regional budgets to universities. The principal objective of this study is to determine if the regional expenditures destined to building a strong scientific capability are related to a greater utilisation of the results of academic scientific research on the part of those companies that generate patentable technology. To draw the regional map of science-technology flows, we assume that the scientific citations in the patent documents filed by the patent applicant are an adequate indicator to represent the use of scientific knowledge by the private sector of industry.[1] The picture we present in this study provides new empirical evidence in respect of the complex science-technology relationships that can be utilised as one additional element in judging the value of future regional action in R&D planning. It should be borne in mind, however, that there exist different forms and mechanisms of transmission between scientific knowledge and industrial technology; in this study we deal with only part of this extensive relationship between science and technology, the part involving the use made of codified scientific knowledge by a company, which is reflected in its patented technology.[2]

This study contributes various novel elements. In the first place, there are no previous empirical research studies carried out for the case of Spain considering how the scientific community influences the transfer of scientific knowledge to the industrial sector. Published research focused on its sectoral characteristics and its regional distribution is practically non-existent.[3] The

[1] In the 1990's, studies such as those of Van Vianen et al. (1990), Grupp and Schmoch (1992), Narin and Olivastro (1992, 1998), Noyons et al. (1994), Narin et al. (1995, 1997), Meyer-Krahmer and Schmoch (1998), among others, have demonstrated that the mean number of scientific references cited in patents is an appropriate indicator for describing science-technology relationships. Recently various analyses with various levels of aggregation have been carried out; these are enabling us to make progress in interpreting the role played by scientific citations in patent documents, for the quantification of science-technology relationships (Meyer 2000a, 2000b, 2000c, 2002; McMillan et al. 2000; Tijssen 2001, 2002; Tijssen et al. 2000; Verbeek et al. 2002). In the section on methodology, we consider the limitations of the procedure.

[2] On a theoretical level, basic science contributes to technological progress through direct factors (it generates useful knowledge) and indirect factors (it provides capacities for resolving problems, forming networks, etc.). Some studies have specifically identified the principal channels and mechanisms of transmission between science and technology. The studies of Lundvall and Borras (1997) and Salter and Martin (2001) are basic references that detail the different paths of interconnection between science and technology. From an empirical point of view, Agrawal (2001) conducted an extensive review of the literature on university-company relationships. The work of Pavitt (1998) contains many references to empirical studies on this aspect of the subject.

[3] The study of Acosta and Coronado (2003) is the only one of which we are aware. In this research science-technology relationships are outlined using the NPC (non patent citation) methodology and it is confirmed that regional differences do exist. But it does not enter into the identification of the factors or explanatory causes that determine those differences.

possible positive effects of expenditure on academic research (and of a favourable scientific environment in general) on the industrial sector, to which we previously referred, have not to date been clearly demonstrated. In other fields, economic geography has paid considerable attention to the contribution to externalities generated by university research. For example, Zucker et al. (2000) find evidence that the entry of companies into the biotechnology sector in the US has been determined by the geographic distribution of high-quality university research in this field. The studies made by Jaffe et al. (1993, 1996, 1998) have utilised data of patents and citations of these patents to quantify the spill-over effects of academic research, and concluded that university research leads to more relevant inventions. In an analysis of the contribution of university-industry interactions to product innovations, MacPherson (2002) concludes that the geographical proximity of the academic research is a factor that helps with the development of production (although this author's data also reveal that non-geographical factors together play a much more relevant role). Empirical research in this area has, in general, detected the presence of local positive externalities of university research for the development of innovations (Feldman and Florida 1994; Varga 1998; Anselin et al. 1997, 2000), although this statement is not universally applicable, being limited to certain sectors and geographic areas.

Secondly, a framework is suggested for identifying, from a microeconomic viewpoint, the features of corporate behaviour in the use of scientific knowledge, and their explanatory causes.

The remainder of the paper is organised as follows. Section 2 provides the basic aspects of the methodology and the hypothesis. The information about the sample and descriptive analysis of the data are outlined in Sect. 3. The results produced by the models specified are presented in Sect. 4. Finally, in Sect. 5 we summarise the principal conclusions.

2. Methodology

2.1. Theory and hypothesis

The empirical tradition for the analysis of the externalities or spillover effects of university research on regional innovation follows the formulation of the "knowledge production function" begun with the work of Griliches (1979), and modified by Jaffe (1989), Feldman and Florida (1994) and Anselin et al. (1997) to include the spatial dimension. Essentially, this involves a neoclassical production function where knowledge is measured by means of a proxy variable (e.g., number of inventions, innovations, etc.) and the inputs incorporate university research, together with other spatial variables, in addition to the classic factors. In this study, we adopt a different approach, since our intention is not to analyse the effects on the regional innovative capacity, but rather the companies' application of scientific knowledge generated in the universities.

The basic reason for studying the transfer of scientific knowledge from university to company, from a microeconomic point of view, starts from the assumption that knowledge in general, and scientific knowledge in particular, will contribute to improving companies' innovative activities and to making them more efficient. In consequence, the corresponding costs to the companies will be reduced and this will ultimately increase the market value of the companies (see Blundell et al. 1999). The relationship between the stock of knowledge (scientific and technological) and the market value of companies has been tested empirically in the framework provided by a specification that Griliches (1981) proposed and Hall et al. (2001) and Bloom and Van Reenen (2000) recently utilised. The factors utilised to explain a company's market value are its stock of knowledge in general, together with other regressors that reflect the internal structure of the company. Knowledge is, however, a difficult input to quantify; in microeconomic models, it has usually been measured by indicators that aggregate scientific and technological knowledge &(R&D expenditure, numbers of scientists and engineers employed, patents weighted by quality level, etc). In this context, our intention is to contribute to the explanation of what are the factors that condition or determine one of the resources of the companies that generate technology: scientific knowledge. To represent this phenomenon adequately, we suggest a function of utilisation of scientific knowledge; this is a variable that will depend both on a company's needs and/or possibilities for adopting this knowledge, and on the opportunities that the firm's environment may offer for its transmission. In other words,

$$SK_i = f(C_i, T_i)$$

Where *SK* represents the application of scientific knowledge by the company *i*, for the development of a particular technology, which is a function of:

– A vector of variables C_i that jointly express the needs and capabilities for the adoption and absorption of scientific knowledge. The need to survive generates differences in the use of the sources of scientific and technological knowledge in favour of some sectors. Various studies have demonstrated that scientific research is relevant for industrial R&D in only a reduced number of industrial activities or sectors, typically in agriculture, chemicals and pharmaceuticals, electronics and precision instruments (Mansfield 1991, 1998; Jaffee 1989; Jaffe, Trajtenberg and Henderson 1993; Audretsch and Feldman 1996; Klevorick et al. 1995; Meyer-Krahmer and Schmoch 1998). In general, those sectors subjected to processes of accelerated technological change and obsolescence usually present closer or stronger relationships with the scientific community than other economic activities.

Moreover, in order for a transfer of scientific knowledge to the industrial sector to take place, the corresponding companies should have a minimum capacity for the absorption of such knowledge. This capacity is usually and primarily related to the learning experiences resulting from undertaking scientific research and technological development activities (Cohen and Levinthal 1989, 1990).[4] Thus R&D activities perform a double function: on the one hand, they contribute to the development of new products and productive processes, and on the other, they strengthen the learning capacity of the company (Schartinger et al. 2002). In this context, Cockburn and Henderson

[4] The concept of absorptive capacity was introduced by Cohen and Levinthal (1989, 1990) and was defined as a company's capabilities for recognising, assimilating and applying new scientific information for its innovations.

(1998) conclude that the actual performance of R&D activities in a company is a necessary but not sufficient condition for stimulating the company's capacity for knowledge absorption: the degree of connexion with the scientific community through the company's participation in research is also important for utilising knowledge spillovers. Lim (2000) finds that companies are capable of exploiting scientific knowledge generated externally without needing to carry out R&D activities internally within the company, but rather by taking advantage of a variety of channels of communication with the scientific community.

– A vector of variables Ti that reflect the territorial opportunities offered by the scientific environment in which a particular company is located. The opportunities for the transmission of scientific knowledge can be defined from the concept of "academic knowledge spillovers"; in other words, as formal and informal movements of new ideas based on science, concepts or technical procedures, from university research to the private sector (Jaffe 1989). A consistent body of literature has built up in support of the hypothesis that physical proximity is important for the flow of knowledge. Jaffe (1989) examined the geographical coincidence of university research with that of research labs within 29 US states for 1972–77, 1979 and 1984 respectively. His results indicate that patents occur in those states where public and private knowledge-generating inputs are the greatest. Even after controlling for industrial R&D, the results indicate that the knowledge generated at universities spill over for higher realized innovative output. Jaffe et al. (1993) examined localization of citation patterns. Taking patent citations as a proxy for knowledge spillovers, they found that US citations to domestic patents were more likely to be domestic. Further, citations were more likely to come from the same state or Standard Metropolitan Statistical Area. Maurseth and Verspagen (1999), in a study across European regions, tested the proximity effect using data on patent citations from the European Patent Office. They found compelling evidence of a localization pattern of patent citations. However, national barriers were important, since patent citations occurred more often between regions belonging to the same country. Verspagen and Schoenmakers (2000) use patent citations between 27 large multinational groups as an indicator of technology spillovers. In order to test for the proximity effect, they measure geographical distance at the level of firms, using data on the location of inventive activities of the firms. Their results generally confirm the proximity effect on spillovers and find significantly negative coefficients on the geographical distance variable. Lastly, Fischer and Varga (2003) examined spillovers of knowledge from universities on patent application activity in 1993. Their sample consisted of firms belonging to one of six technology classes in political units in Austria. Employing a spatial econometric approach, the authors find evidence of spillovers across regions, but that this is linked to a spatial decay effect. Their empirical results confirm the presence of geographically mediated knowledge spillovers from university and show that these transcend the geographic scale of the political district. This result is in line with the previous studies conducted by Anselin et al. (1997, 2000) despite differences in research design and context.

On the basis of these results, one would think that a company's location in a territorial environment with a well-established university presence increases the possibilities for the company to access relevant new scientific knowledge more rapidly, in comparison with other companies located elsewhere, because geographic proximity reduces communication costs, especially for face-to-face contacts (Fritsch and Schwirten 1999; Goe et al. 2000). Various analyses with aggregated data support this hypothesis, although not conclusively (Anselin et al. 1997, 2000). The basic hypotheses that we wish to test in this study concern the possible influence of the regional scientific community or environment on the use of science by companies; we can specify these hypotheses as follows:

Basic hypothesis:

H1: *Science-technology relationships depend on the scientific environment in which companies undertake their industrial activity.*

This general hypothesis can be broken down into four more specific hypotheses involving indicators of the resources and results of scientific research that give some measure of the scientific environment:

In respect of the scientific resources available:

H2: *Larger allocations of regional expenditure to university research contribute to increasing the use made of science by regional industry.*
H3: *Larger allocations of human resources dedicated to local university research contribute to increasing the use made of science by regional industry.*

In respect of the results of scientific research

H4: *The generation of scientific output by local universities contributes to increasing the use made of science by regional industry.*
H5: *The generation of competent human resources (numbers of trained scientists and engineers) by local universities contributes to increasing the use made of science by regional industry.*

2.2. The variables

To quantify the variable *SK*, we utilise the scientific citations in patent documents as valid indicators reflecting the application of scientific knowledge in industry. We make the same assumption as in other recent studies (Meyer 2000a, 2000b, 2000c, 2002; McMillan et al. 2000; Tijssen 2001, 2002; Tijssen et al. 2000; Verbeek et al. 2002) that the scientific citations in patent documents (NPC) are a proxy variable to represent a particular unit or item of scientific knowledge that is required or considered useful for the development of patented technology. Patented inventions generally incorporate public and private knowledge in different proportions, and this materialises in references to or citations of other patents and scientific literature. In patent documents, as occurs in scientific articles, it is usual to provide references or citations, the objective of which is to describe the antecedents or "state of the art" prior to the invention. The antecedents or state of the art include not only other patents that have been utilised as support for the invention, but also bibliographical references to scientific literature and technical publications. These citations provide some indications of the potential contribution of the published research to the inventions patented. But the essential question to confer validity to the procedure followed in this study is: Do the scientific citations in patent documents genuinely reflect the use made of scientific knowledge by the industrial sector? The following arguments will help to answer this question:[5]

- The references included in a patent are less likely to be redundant or superfluous than those incorporated in a scientific article, due to the control that is exercised over patents and to their legal consequences (Collins and Wyatt 1988; Verbeek et al. 2002).[6]
- The empirical analyses conducted in various studies (Grupp and Schmoch 1992; Schmoch 1993; Narin and Olivastro 1998; Meyer 2000a 2000b) identify a number of reasons why examiners, applicants or inventors incorporate NPCs in patents or do so with different intensities or frequencies, and not all are related to possible interrelationships with science.[7]
- The studies in aggregate seem to confirm that NPCs measure the intensity of the science supporting the innovations patented; however, a more diffuse panorama appears when one tries to determine the type of relationship between research published and inventions patented (Tijssen et al. 2000; Tijssen 2002). Narin et al. (1997) are relatively optimistic in respect of the NPC methodology for measuring the relationships of dependence of technology on academic research (they utilise this conclusion to argue that "public science is a driving force behind high technology"). Other authors are less optimistic when it comes to describing science-technology links and prefer to speak of interactions (Schmoch 1993). Therefore, "it does not seem appropriate to use the linear science-push model to interpret patent citation data" Meyer (2000a), and also "One should refer to science and technology *interplay* rather than speak of science-dependence in the context of patent citations" (Meyer 2000c). This lack of clear causality is also noted by Tijssen who refers to "the questionable validity of these citations as causal measures of knowledge flows from the science base to the technology

[5] For more details, a review of the NPC procedure is given in Acosta and Coronado (2003), and the bibliography indicated can also be consulted.

[6] For an extensive empirical analysis of the differences between the citations included in a patent and those incorporated in a scientific article, see Meyer (2000b).

[7] For example, the limited availability of patents in particular technological fields in consequence of the rapid advance of certain technologies and the consequent time-lag in the publication of the patent documents, the legal context of patents (their obligatory nature, and the responsibility of including discussion of the prior art, utility, novelty, etc. of the invention), the social nature of the process (involvement of various actors - inventor, examiner, attorney, etc.) are integral elements in the development of the patent and exert influence on its final form. In addition, there are differences in national practices (different patent offices have different methods of work; it is well known that, in USPTO patents, the frequency of citations is higher in comparison with EPO patents).

domain'' (Tijssen 2002). In summary, various validation studies seem to have lowered the degree of optimism expressed in the initial studies of Narin and his co-workers when interpreting the results obtained from the study of NPCs, and it seems that it would be more appropriate to speak of interactions, links or science-technology flows, in place of relationships of causality, of dependence or of science driving technology, as postulated by a linear model.

– In the particular case of Spanish patents, we can add to these preceding reasons several others that could affect the relative propensity to cite. Normally the patents applied for by universities cite proportionately much more frequently than those applied for by companies or private individuals. Apart from the greater knowledge of scientific literature presumably available in a university, another possible reason could be the differences in the objectives of the various kinds of applicant. The businessman wishes to secure protection for the commercial exploitation - in a monopolistic position - of a product or a productive process; for the university applicant, the patent is an academic merit, although it may also be exploited commercially. Furthermore, the patent applications that are presented in the OEPM (Spanish Patent Office) may be completed by professionals or by the inventors themselves. In this latter case, more deficiencies are apparent when it comes to collecting and documenting the antecedents and even in the actual description of the invention (Carlos Velasco, personal communication).[8]

The conclusion after presenting these arguments can be summarised as follows: the scientific citations in patent documents are useful indicators for reflecting the relationships between the scientific and industrial fields, but they only constitute a partial perspective of the complex science-technology relationships. Scientific citations only reflect one part of the contribution of science to technology, above all, the part of codified knowledge that is utilised as a source of ideas, analytical methods and data: other interesting sources of tacit knowledge are not visible by means of this methodology.

Further, our work presents two fundamental differences with respect to the empirical studies that have been based on the analysis of scientific citations for analysing the behaviour of science-technology relationships:

– First, unlike the approach that is becoming habitual, our analysis has been made utilising the domestic patents (NLP), rather than European (EPO), international (PCT) or American (USPTO) patents. The reason for this choice is the actual delimitation or definition of scope of this study; the use of national patents widens the information: domestic patents are the most numerous, and distortions are not introduced (high quality technology that pursues protection by other routes is not left on the margin, since most of the patents of Spanish companies that follow an European, international or American route have previously been applied for through the national NLP route). To this it must be added that, in many Spanish regions (above all, in

[8] C. Velasco is Chief Examiner of Chemical Patents of the Spanish Office of Patents and Trade Marks (OEPM).

the less developed regions) the EPO, PCT or UPSTO patents are not very representative of the technology that is developed in these zones.

– Second, other studies that have employed this indicator are based on the citations included by the examiner of the patent; these do not always coincide with those made by the applicant. In our case, the citations collected and classified are those included in the complete text; in other words, they are the citations made by the inventor, and not those added by the examiner. In our opinion, the inventor's citations better represent science-technology relationships and the industrial use of scientific research, since the citations of the examiner have been added for other purposes.[9]

Accepting the considerations and limitations indicated in the previous paragraph, the NPC citations made by each company in the text of their patent document are the dependent variable that we shall attempt to explain by the following explanatory factors that, as mentioned earlier, are divided into two groups:

2.2.1. Needs and capacities for the absorption and adoption of scientific knowledge

In order to reflect the need and ability to absorb scientific knowledge empirically, the literatura in this field initiated by Cohen and Levinthal (1989, 1990) has emphasised the performance of R&D activities as a fundamental, but not unique instrument for strengthening the capacity of a company to absorb scientific knowledge. In our model we reflect the potential of companies for undertaking R&D activities (absorption capacity) by means of four indicators: a) The activity sector in which the company operates; we will employ two binary variables (Gji, $j = 1$ to 4) which take account of the sector (technological group) involved in the development of new technologies. b) Its capacity for utilising technological knowledge. We reflect this indicator using a variable (TK) defined as the average number of patent citations drawn from the patent documents of company i. This variable takes account of any other sources of (technological) knowledge employed by the company, other than scientific knowledge. The variable TK quantifies the induced benefits for an inventor derived from inventions made by others. c) Its market power or technological leadership. With the object of capturing the influence obtained from a dominant position in the market or from technological leadership, we include in our model a variable (L) defined as the number of patents of company i in sector j, relativised by the total number of patents in the sectoral "block" (of which there are 5) in which sector j is included. d) Its degree of

[9] The citations may have been made by the applicants for the patent or by the examiners in the process of evaluation of the patent (in the body of the patent, or on the front page, respectively). The examiners are more inclined to take NPCs when the invention is not sufficiently documented with previous patents (Grupp and Schmoch 1992; Tijseen et al. 2000). In a case study, Meyer (2000b) illustrates how the citations of the examiners are related to those of the applicants, and show different types of behaviour; and although this study does not allow a global conclusion to be drawn, it does demonstrate that the examiners do not include all the citations made by the applicants.

diversification, reflected by a variable (D) which takes account of the number of sectors in which company i files patents.

2.2.2. Opportunities of the environment for the transmission of scientific knowledge

With the object of quantifying the variables related to the scientific environment, various indicators associated with resources and scientific results have been utilised in the empiric literature. For example, Audrestch and Feldman (1996) find a positive relationship between "local university research funding" and "local industry value-added" at the state level. Their results indicate the relative economic importance of new knowledge to the location and concentration of industrial production. Zucker et al. (2000) relate the input "number of local research stars" to the output "number of new local biotech firms" and examine the variance in this relationship across geographic space at the economic region level. They find that the number of local stars and their collaborators is a strong predictor of the geographic distribution of biotech firms in 1990. Branstetter (2001) identifies a relationship between "scientific publications from the University of California" and "patents that cite those papers", also at the state level.

In order to bring in these factors, in this first block we incorporate into the function $Sk_I = f(\cdot)$ a set of variables that represent the regional scientific environment: a) two indicators of inputs, and b) two of outputs.

a) Inputs: As the fundamental variable, we have chosen the regionalised expenditure allocated to university research, per inhabitant, this being the variable that quantifies the efforts made to build up the regional scientific research capabilities. In addition, on the input side, we take the number of university researchers (per thousand inhabitants).
b) Outputs: Two indicators have been taken; one to represent the situation of the scientific environment by means of the number of students completing science and technology courses – excluding social sciences, law and humanities- per thousand inhabitant, and another that shows the results of scientific research, quantified by the number of doctoral theses in fields of science and technology. All the variables have been adjusted for the number of inhabitants.

Lastly, in a microeconomic study of North American biotechnology companies, Audretsch and Stephan (1996) find that physical proximity between academic research and biotechnology companies is positively related to the existence of collaboration between the two fields. They conclude, however, that this evidence is not absolute, and that relationships can be maintained effectively over long distances by companies prepared to invest in collaboration with the academic community (Audretsch and Stephan 1996). In consequence, we are presented with two possible types of spillover effect: horizontal effects, generated by the location of the company in a particular scientific environment, and vertical effects, derived from the company's collaborative relationships with the academic community. We have accounted for possible vertical spillover effects of scientific knowledge by means of a dichotomous variable that indicates whether or not the patenting company has collaborated with a university or public research institution in the development of the patent. The source of the variable is the Spanish Patent Office. In Table 1, each of these variables and its definition are given.

3. Data

The primary information has been obtained as follows. Our sample is formed by 1,643 patents (registered by 1,129 companies) that represent the totality of patents presented by the national route in Spain in the period 1998–2001, both years inclusive. Of the 1,643 patents, 9.92% (163 patents of 79 companies) contain NPC scientific citations (references to scientific literature, text

Table 1. Definition of explanatory variables

Name	Definition
Needs and capacity for the absorption and adoption of scientific knowledge	
Gji (j = 1 to 4)	Binary variables which take account of the technological sector j (Group) in which company i files patents (see sectors in Table 1). Gji has a value of 1 if company i is in sector j (1 to 4), otherwise, the value is 0. The base category is Gji(j = 5).
Di	Variable which takes account of the number of sectors in which company i files patents. It represents the technological diversification of the company.
Lij	Variable defined as the number of patents of company i in sector j, relativised by the total number of patents in the sectorial 'block' (of which there are 5) in which sector j is included. The relativisation has been based on the number of patents per technological area (see Table 2) and not per sector, in order to avoid the bias which would be generated in the event that one sector presents only very few patent documents. It reflects the leadership of company i in sector j.
Tki	Variable defined as the average number of patent citations drawn from the patent documents of company i. This variable takes account of any other sources of (technological) knowledge employed by the company, other than scientific knowledge.
Vertical spillovers	
Ci	Binary variable which has a value of 1 if company i has collaborated with a university or public institution in submitting the patent; otherwise, the value is 0. This variable takes account of vertical spillovers of scientific knowledge.
Horizontal spillovers: Regional Opportunities for the transfer of scientific knowledge	
Sri	Variable indicating the expenditure on university research of the region r (relativised by the number of inhabitants) in which the company i is geographically located.
Ir	Variable indicating the number of researchers in the university of the region r (relativised by the number of inhabitants) in which the company i is geographically located.
Er	Variable indicating the number of university graduates employed in the region r (relativised by the number of inhabitants) in which the company i is geographically located.
Or	Variable indicating the results of scientific research quantified by the number of doctoral theses presented... (relativised by the number of inhabitants) of the region r in which the company i is located.

Source: Own elaboration

books and other citations), and 90.08% of patents of the sample do not incorporate scientific citations (1,480 patents of 1,050 companies).

The NPC references recorded for these 163 patents amounted to 1,427 citations, of which 969 related to scientific journals included in the Institute for Scientific Information (ISI) Current Contents; consequently, 67.90% of all the NPC references correspond to what is generally accepted as "quality scientific research". The patents of each company were classified according to sufficiently specific criteria capable of distinguishing between five technological areas and thirty subfields based on the International Patent Classification (IPC).[10] (See Table 2).

The units of observation in the models calculated are companies. The number of scientific citations of each company (the endogenous counting variable) was obtained as follows: a) If the company has only one patent, the count variable is simply the number of scientific citations incorporated in the text of the patent. b) If the company has two or more patents in the same sector, the count variable is the average of the number of citations in each patent (rounded to the nearest whole number). c) If the company operates in two sectors, for example, that is, it is technologically diversified with patents in those two sectors, two observations are taken and the procedure is like that in a) when the company has only one patent in that sector, or in b) if the company has more than one patent. Taking this counting method into account, the number of observations in our models totalled 1,139. In respect of the territorialisation of the data, this was taken from the physical location of the company. Table 2 summarises the basic data.

The analysis of these data give us some interesting patterns in the regional diversity of science-technology flows. The main characteristics of the picture revealed by these data are the following:

– *Territorial concentration of science-technology flows*. Catalonia, Madrid and Navarre are, in this order, the three autonomous regions that account for the largest flows, with 69.53% of the NPC citations, and 71.61% of the scientific citations (ISI). The concentration of citations is greater than that of the number of patents (these same three regions account for 62.93% of the patents), which leads one to think that the concentration of science-technology flows is even more polarised territorially than the technology itself. With the object of avoiding the bias introduced by the aggregation of all the different sectors of industry, some indicators of regional concentration have been calculated for the sectors that incorporate most scientific citations in patent documents, and the level of concentration of patents has been obtained for those regions that account for most of these flows. It should be noted that these indicators again demonstrate a strong regional concentration in science-technology relationships. However, a relationship between concentration of technology (numbers of patents in a sector) and science-technology flows (scientific citations in the patents of the same sector) does not necessarily exist; for example, in sector 11: "Pharmaceutical and cosmetic products", the region accounting for the most citations is Navarre, with 45.86% (ISI) of the national total, while in volume of

[10] This classification was devised jointly by FhG-ISI, the French Patent Office (INPI), and the Observatoire des Sciences et des Techniques (OST).

Table 2. Patents, companies, scientific citations (NPC) and patent citations (PC) 1998–2001 sectorial classification

Technology sector	Patents		Companies			NPC citations				PC citations	
	N°	Spain = 100	N°	Spain = 100	N° PAT/N° COMP	Total N°	Total Spain = 100	ISI Spain = 100	%ISI / TOTAL	N°	SPAIN = 100
I. Electrical engineering											
1. Electrical machinery and equipment, electrical power	139	8.46	68	6.02	2.04	4	0.28	0.21	50.00	87	5.31
2. Audiovisual technology	37	2.25	27	2.39	1.37	0	0.00	0	–	10	0.61
3. Telecommunications	57	3.47	32	2.83	1.78	8	0.56	0.1	12.50	12	0.73
4. Information technology	13	0.79	12	1.06	1.08	0	0.00	0	–	18	1.1
5. Semiconductors	4	0.24	3	0.27	1.33	0	0.00	0	–	1	0.06
II. Instruments											
6. Optical	7	0.43	3	0.27	2.33	0	0.00	0	–	0	0
7. Analysis, measurement and control technology	127	7.73	83	7.35	1.53	12	0.84	0.62	50.00	111	6.77
8. Medical technology	54	3.29	42	3.72	1.29	11	0.77	0.52	45.45	29	1.77
III. Chemicals and pharmaceuticals											
9. Fine Organic Chemicals	80	4.87	36	3.19	2.22	500	35.04	37.87	73.40	250	15.25
10. Macromolecular chemicals, polymers	10	0.61	9	0.8	1.11	8	0.56	0.31	37.50	19	1.16
11. Parmaceutical and Cosmetic products	58	3.53	33	2.92	1.76	221	15.49	16.2	71.04	166	10.13
12. Biotechnology	26	1.58	10	0.89	2.6	466	32.66	34.16	71.03	91	5.55
13. Materials, metallurgy	42	2.56	32	2.83	1.31	35	2.45	0.72	20.00	75	4.58
14. Chemicals for agriculture and food industries	56	3.41	51	4.52	1.1	67	4.70	4.85	70.15	70	4.27
15. Chemical and petroleum industries, chemical treatment of basic materials	34	2.07	25	2.21	1.36	58	4.06	2.99	50.00	44	2.68
IV. Process engineering special equipment							0.00				
16. Chemical engineering	40	2.43	34	3.01	1.18	6	0.42	0.1	16.67	21	1.28
17. Paint and surface coatings technology	19	1.16	19	1.68	1	9	0.63	0.62	66.67	25	1.53
18. Processing of materials, textiles, paper.	74	4.5	58	5.14	1.28	0	0.00	0	–	35	2.14

Table 2. (contd.)

Technology sector	Patents		Companies			NPC citations				PC citations	
	N°	Spain = 100	N°	Spain = 100	N° PAT/N° COMP	Total N°	Total Spain = 100	ISI Spain = 100	%ISI / TOTAL	N°	SPAIN = 100
19. Thermal processes and equipment	37	2.25	20	1.77	1.85	0	0.00	0	–	31	1.89
20. Environmental technology	26	1.58	24	2.13	1.08	5	0.35	0.31	60.00	38	2.32
V. Mechanical engineering machinery											
21. Machine tools	52	3.16	38	3.37	1.37	0	0.00	0	–	42	2.56
22. Motors. pumps and turbines	23	1.4	16	1.42	1.44	3	0.21	0	0.00	20	1.22
23. Mechanical items	46	2.8	29	2.57	1.59	0	0.00	0	–	28	1.71
24. Mechanical handling, printing	138	8.4	104	9.21	1.33	0	0.00	0	–	91	5.55
25. Machinery and equipt. for agric. and food processing	70	4.26	58	5.14	1.21	9	0.63	0.41	44.44	74	4.51
26. Transport	79	4.81	53	4.69	1.49	0	0.00	0	–	66	4.03
27. Nuclear power engineering	2	0.12	2	0.18	1	0	0.00	0	–	15	0.92
28. Space and defence technology	10	0.61	7	0.62	1.43	0	0.00	0	–	18	1.1
29. Capital and consumer goods	144	8.76	100	8.86	1.44	5	0.35	0	0.00	75	4.58
30. Civil engineering, construction and mining	139	8.46	101	8.95	1.38	0	0	0	–	77	4.7
Total	1.643	100	1.129	100	1.46	1.427	100.00	100	67.90	1.639	100

Source: O.E.P.M. and own elaboration.

Table 3. Descriptive statistics

	Mean	Std.dev.	Maximum	Minimum	Source
D	1.356	0.838	7	0	O.E.P.M.
L	0.004	0.005	0.088	0.001	O.E.P.M.
TK	0.891	2.221	21	0	O.E.P.M.
C	0.015	0.124	1	0	O.E.P.M.
G1	0.127	0.333	1	0	Table 1
G2	0.114	0.318	1	0	Table 1
G3	0.171	0.376	1	0	Table 1
G4	0.137	0.344	1	0	Table 1
S	6.951	1.421	8.489	2.736	I.N.E.
I	0.954	0.256	1.988	0.255	I.N.E.
E	2.214	0.660	4.860	0.357	I.N.E.
O	0.112	0.039	0.207	0.003	I.N.E.

O.E.P.M.: Spanish patent office. I.N.E. Instituto Nacional de Estadística.

patents, this sector represents only 5.17% of the national total; if we sum the two regions that cite most, Navarre and Catalonia, together they account for 82.17% in number of ISI citations, and 62.07% in number of patents. Something similar occurs in sector 14: Agriculture and food chemistry, where Castille and Leon is the region that incorporates most scientific citations (27.66%), whereas the national participation of this sector in number of patents is only 5.36% the two regions with the most citations are Castille and Leon and Andalusia, together accounting for 55.32% of the national total of scientific citations, and 12.50% of the total national patents. These data suggest that the overall tendency of concentration in these flows is conditioned by the relative weight of technological sectors with high propensities to cite scientific literature and of traditional sectors where advances are basically supported by previously patented technology.

- Regional behaviour in the business concentration of the science-technology flows. The ratio of distribution of the number of citations in relation to the distribution of the number of companies gives us a general picture of the degree of regional concentration of science-technology flows in the companies located in each region. Madrid, Andalusia and Navarre present indicators higher than unity, this denoting a greater concentration of flows in relation to the number of companies that seek patents. It is interesting to see that Navarre presents a figure almost three times the national average: this region has only 3.52% of the companies that patent, but accounts for 10.43% of the NPC citations and 10.84% of the ISI citations. The opposite pattern of behaviour is presented by the Region of Valencia, which has 11.44% of the total companies that patent, but accounts for only 3.64% of the NPC citations.
- Application of "quality" science. To study such application we have calculated the ratio of the number of scientific ISI citations, to the total number of NPC citations. The Basque Country, Navarre, Madrid and Catalonia, in this order, are the autonomous regions that utilise "quality" science in greater proportion than the rest (with ratios of ISI citations to NPC citations of around 70%).

– *Scientific base and technological base of the inventions patented.* Finally, the ratio of the "national proportion of NP citations (patent citations) of each region to the proportion of NPC citations" has been obtained. Navarre, Madrid and Andalusia have values lower than unity; this means that, in relative terms (in relation to the Spanish average), for the totality of sectors, these regions are more intensive in scientific than in technological knowledge; while in other autonomous regions such as Catalonia, the Basque Country and Valencia, where technological change is supported more by the development of previously patented technology (citations of patents), relatively less use is made of scientific knowledge.

In Table 3 we give the basic data of the explanatory variables that will be incorporated in the models.

4. Econometric specification and results

4.1. Econometric specification

The empirical treatment of the function of utilisation of scientific knowledge SKi = f(Gji, Di, Lij, TKi, Ci, Sri, Ir, Er, Or) suggests various possibilities:

a) In the first place, the nature of the data points to the formulation and estimation of a count model to quantify the intensity in the use of the knowledge, by using the number of citations (Poisson or negative binomial). The application of the Poisson model requires equality of means and variance, a requirement that cannot always be met in practice. If the data show overdispersion, the standard errors of the Poisson model will be biased to the low end, giving spurious high values for the t statistics (Cameron and Trivedi 1986, p. 31). The most common formulation for taking into account the overdispersion of data is the negative binomial model (NB2, in the terminology of Cameron and Trivedi 1986). This assumes that the variance is a quadratic function of the mean (the proposal of the density function, the logarithmic likelihood function, the first order conditions, etc, is comprehensively detailed in Cameron and Trivedi 1998, p. 71 and following).
b) In the second place, from observation of the data, with a large number of zeros in the sample, one would tend to think that the process generating the data is possibly formed by two regimes: One that traces the access to scientific knowledge and another that defines its intensity. Two categories of modelling can be utilised for the excess of zeros. One of these consists of estimating hurdle models suggested by Mullahy (1986). In this formulation, a first distribution determines if the count variable has zero or positive results. If the result is positive, a second distribution, truncated at zero, determines the result. In our case, we assume that the underlying distribution for the first step is logistic, and we model the decision by a logit model. A Poisson or negative binomial distribution governs the second step (the details of the density function, the likelihood function and methods of estimation can be found either in Mullahy 1986, in Crepon and Duguet 1997, or in Cameron and Trivedi 1998). A more general formulation, introduced by Lambert (1992) and analysed in depth by Greene (1994) is the Poisson model with probability of zero increased (zero

inflated Poisson model ZIP), where a zero result can be originated either by a binary process, or by a Poisson process. In other words, the fact that a sampling value is zero can be produced either because the company has no opportunities for access to scientific knowledge (the geographical circumstances prevent this), or simply because, despite having the opportunities and capacities, it has no need for the knowledge because its patented technology has been based more on previously existing technology (technological knowledge) than on scientific research (scientific knowledge). Both models, the hurdle and the ZIP model, allow two sources of overdispersion: one that permits a number of extra zeros, and another introduced by the individual heterogeneity of the set with positive values (see details in Cameron and Trivedi 1998, p. 123 and following).

4.2. Results

With the object of producing more robust results, the two categories of modelling formulated in the econometric specification part of this paper have been estimated. For each category of model, six specifications have been estimated: a first model containing no variables of regional scientific environment, a second model that incorporates jointly all the regional variables, and four more models that include the variables of regional scientific environment separately. The reason for treating the variables separately is that there is a strong linear relationship between them, and this could cause false individual results when they are considered together in the same model, and even unjustified changes of sign. The hypothesis H1 has been tested by applying a LR test of the model with all the variables reflecting the scientific environment, against the model that excludes these variables (in all the specifications estimated, the test LR = -2[lnL*Model* I-ln*Model*II] is calculated, and compared with a chi squared with four degrees of freedom). The hypotheses H2, H3, H4 and H5 have been tested by analysis of individual significance on the coefficients of each variable. In the following, we comment on the results of the estimations and statistical tests:

a) Poisson/NB models. The models estimated using a Poisson or NB distribution consider the number of scientific citations in the patents to be a count variable and, consequently, are able to account for the intensity of scientific knowledge use. However, if the Poisson model presents overdispersion (the conditional variance exceeds the conditional mean), the standard errors of the P model are biased towards the low end, which results in high values of the t statistic.[11] In the case of the Poisson model, the application of the overdispersion test developed by Cameron and

[11]The overdispersion tests are tests of equality of means-variance, imposed by the Poisson specification, against the alternative case that the variance exceeds the mean. These tests, such as the LR, Wald or LM tests that are asymptotically equivalent, are easy to apply. However, the usual critical values of the LR and Wald tests cannot be used and adjustments are needed because the null hypothesis alpha = 0 falls on the boundary of the negative binomial, which does not allow overdispersion (Cameron and Trivedi 1998, p. 159).

Table 4. Results of Poisson models

	Model I			Model II			Model III			Model IV			Model V			Model VI		
	Coeff.	Std.err.	Robust	Coeff.	Std.err.	Robust	Coeff.	Std.err.	Robust	Coeff.	Std.err.	Robust	Coeff.	Std.err.	Robust	Coeff.	Std.err.	Robust
ONE	−4.846*	0.383	0.643	−6.018 *	0.476	0.994	−5.764 +	0.431	0.932	−6.166 *	0.402	0.800	−5.939 *	0.399	0.789	−5.530 *	0.399	0.818
L	3.899 *	0.553	1.357	3.237 *	0.607	1.861	3.590 *	0.566	1.382	3.159**	0.603	1.769	3.063**	0.606	1.735	3.333 *	0.586	1.507
D	0.270 *	0.033	0.095	0.243 *	0.034	0.103	0.262 *	0.034	0.092	0.249 *	0.034	0.097	0.263 *	0.035	0.096	0.284 *	0.035	0.096
C	1.249 *	0.098	0.288	1.440 *	0.101	0.292	1.267 *	0.098	0.284	1.420 *	0.100	0.290	1.377 *	0.099	0.285	1.276 *	0.098	0.279
TK	0.036	0.011	0.036	0.042	0.011	0.035	0.032	0.011	0.037	0.042	0.011	0.035	0.040	0.011	0.035	0.035	0.011	0.037
G1	1.722**	0.476	0.974	1.805**	0.476	0.968	1.673**	0.476	0.971	1.787**	0.476	0.971	1.755**	0.476	0.971	1.685**	0.476	0.970
G2	2.490 *	0.433	0.823	2.508 *	0.432	0.811	2.440 *	0.433	0.818	2.491 *	0.432	0.809	2.471 *	0.432	0.809	2.444 *	0.432	0.814
G3	5.303 *	0.381	0.646	5.345 *	0.381	0.643	5.282 *	0.381	0.645	5.334 *	0.381	0.642	5.322 *	0.381	0.642	5.289 *	0.381	0.644
G4	2.222 *	0.436	0.797	2.246 *	0.437	0.799	2.238 *	0.436	0.794	2.241 *	0.437	0.797	2.253 *	0.437	0.795	2.236 *	0.436	0.793
S				−0.004	0.055	0.171	0.136	0.029	0.094									
I				0.982	0.595	1.002				1.334 *	0.114	0.449						
E				0.273	0.299	0.556							0.476 *	0.044	0.185			
O				−3.435	2.801	8.335										5.960	0.943	4.161
−lnL		1235.93			1178.57			1224.04			1179.9			1186.91			1215.53	
P^2			0.625			0.657			0.636			0.658			0.655			0.645
D^2			0.516			0.542			0.522			0.542			0.538			0.523
Overd(1)			4.007			3.704			4.057			3.762			3.821			4.049
LR (2)						114.72 *												
N obs.		1139			1139			1139			1139			1139			1139	

* Sign. 5%; ** Sign.10% (for robust estimations).
(1) Overdispersion test by Cameron and Trevidi (1990).
(2) LR = −2(lnL*Model*I−ln*Model*II)

Trivedi (1990) (based on a t statistic) suggests the presence of over-dispersion in the data (see overdispersion in Table 4).[12]

The assumption of equality of means and variance does not seem reasonable. A feasible option for avoiding the bias towards the low end of the standard errors of the Poisson model is to apply the correction of Eicker-White[13] to obtain robust values (these values are given in the "robust" column for each model in Table 4). Individually the variables of regional environment in Model II do not achieve significance, in consequence of the colinearity between them; therefore, the hypothesis H1 has been tested by applying the test LR = -2(lnL*Model*I-ln*Model*II) of joint significance on the regional environment variables *S*, *I*, *E* and *O* of model II, against model I that does not include these variables. The value of the LR statistic is 114.7. When this is compared with a chi squared with 4 degrees of freedom (9.49), it tells us that jointly the variables are relevant; therefore, the robust estimations of Poisson favour the hypothesis H1.

The introduction of these variables individually (Models III, IV, V and VI) leads us to reject hypotheses H2 and H5; this means that the parameters of the variables *I* and *E* are statistically significant. Therefore, taking into account the results of the robust estimations of the Poisson models, the characteristics of the regional scientific environment are relevant in explaining the utilisation of science by the companies, but it is the human resources (I: number of university researchers, and E: number of science, technology and engineering graduates) that are the explanatory factors, and not the expenditure (S) nor the scientific research results (O). Considering the other explanatory variables, it can be observed that the technological leadership of the company (L), its degree of diversification (D), its collaboration with universities (C) and the type of technological activity undertaken (G2, G3 and G4) are also relevant factors for explaining company behaviour regarding the utilisation of scientific knowledge (Table 4).

As an alternative to the robust estimations of the Poisson model, Table 5 gives the results of the more common formulation for taking account of the overdispersion of the data (negative binomial or NB2, in the terminology of Cameron and Trivedi 1986). The coefficient of overdispersion (alpha) is statistically significant in all the cases, showing the adequacy of this type of model. The LR test of joint significance on the variables of regional environment *S*, *I*, E and O of model II against model I tells us that they are jointly relevant; therefore the estimations of the negative binomial model favour hypothesis H1. The introduction of all these variables individually (in Models III, IV, V and VI of Table 5) shows the expected signs and their parameters are statistically significant for the variables reflecting both the resources and the results. In consequence, the estimations of the NB models also confirm the hypotheses H2, H3, H4 and H5. With respect to the rest of variables, all are

[12] The same results are obtained by applying other customary tests of overdispersion that require the estimation of the negative binomial model: the t tests on the alpha parameter of all the NB models indicate that they are significant, and if we compute the test $LR = 2(\ln L_{NB}\text{-}\ln L_P)$, overdispersion is confirmed, the same as in the Poisson model.

[13] Eicker (1967) and White (1982).

Table 5. Results of negative binomial models

	Model I		Model II		Model III		Model IV		Model V		Model VI	
	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.
ONE	−5.101	0.555	−8.372*	1.416	−6.972*	0.998	−7.212*	0.950	−6.692*	0.832	−5.968*	0.705
L	0.183	3.334	−0.336	3.285	0.069	3.407	−0.315	3.283	−0.265	3.339	0.007	3.422
D	0.453*	0.201	0.332**	0.193	0.382**	0.197	0.344**	0.192	0.346**	0.192	0.388*	0.196
C	1.511**	0.773	1.933*	0.774	1.543*	0.765	1.918*	0.773	1.820*	0.774	1.483*	0.770
TK	0.148*	0.048	0.163*	0.051	0.138*	0.048	0.158*	0.050	0.147*	0.049	0.135*	0.048
G1	1.944*	0.578	2.052*	0.616	1.695*	0.582	2.140*	0.603	1.981*	0.588	1.764*	0.580
G2	2.624*	0.556	2.687*	0.585	2.561*	0.556	2.820*	0.584	2.794*	0.576	2.693*	0.561
G3	5.184*	0.500	5.326*	0.530	5.203*	0.502	5.387*	0.539	5.358*	0.528	5.252*	0.505
G4	1.571*	0.573	1.743*	0.604	1.504*	0.576	1.829*	0.605	1.744*	0.594	1.565*	0.575
S			0.228	0.211	0.282*	0.119						
I			3.933	3.294			2.057*	0.655				
E			−0.976	1.647					0.691*	0.234		
O			−0.082	12.407							8.241*	3.823
Alpha	9.636*	1.645	8.934*	1.518	9.334*	1.585	9.066*	1.533	9.199*	1.549	9.506*	1.604
−lnL	502.55		496.14		499.74		496.80		497.60		500.12	
LR (1)			12.82 *									
N obs.	1139		1139		1139		1139		1139		1139	

* Sign. 5%; ** Sign.10%.
(1) LR = −2(lnL*Model*I−ln*Model*II)

relevant with the exception of variable L (technological leadership); this parameter is not statistically significant in any of the models in Table 5.
b) ZIP and ZINB models. When the data generating process is characterised by a dual regime, the NB distribution may indicate a spurious overdisperson. Consequently, ZIP or ZINB specification analysis would be appropriate to determine the possibility of mixed distributions. With two regimes being present, we can assume that one will be determined by the needs and opportunities for access to scientific knowledge, and the other will determine its intensity. In other words:

– The necessities are included through the binary sectoral variables ($G1$, $G2$, $G3$ and $G4$), the opportunities are included through the variables reflecting the scientific environment (S, I, E and O) and collaboration with the university in the development of the patent (C). The formulation incorporated in the ZIP or ZINB specification to include these needs and opportunities is of the binary type (logit), which determines either access or no access to scientific knowledge.
– The intensity in the application of scientific knowledge, may be reflected by company variables, and by determinants of the company competence (D, L, TK), as well as by all the preceding variables that not only allow the companies access to a particular type of scientific knowledge (variables of environment) or determine their need, but can also influence the intensity of use. The specifications to cover this regime are of the Poisson (ZIP) or negative binomial (ZINB) type.

The estimations of both specifications -ZIP and ZINB- are favourable to hypothesis H1. For the ZIP model, the LR statistic is 12.24 (Table 6), and for the joint significance of the parameters of the scientific environment, it can be observed that jointly the variables S, I, E, and O are relevant. For the ZINB specifications, this statistic is 13.88 (Table 7). The introduction of these variables individually into both groups of model shows the individual significance of all of them; therefore these models are also favourable to hypotheses H2, H3, and H4.[14] It can be observed, however, that the variable S (university expenditure on R&D) is only relevant in the splitting, in other words, in the part of the ZIP that determines the access to scientific knowledge, but not in the Poisson that determines its intensity (Table 6). In Table 7 this same circumstance can be appreciated for all the variables of the scientific environment (they are statistically significant in the *logit* part of the ZINB, which determines the access to scientific knowledge, but are not so in the negative binomial, which conditions its intensity).

For both the ZIP and the ZINB models, the Vuong (1989) statistic selects the model with two regimes, against the Poisson or NB models.

The joint analysis of the results obtained in the testing of the hypotheses is summarised in Table 8, showing the hypotheses, the variables that represent each, and the effects (positive/negative) in each model.

[14] The negative signs of the Logit specification of the ZIP and ZINB occur because the value 1 is assigned when $y = 0$, and value 0 when $y > 0$. In consequence, a negative sign represents a relationship in the same direction between the utilisation of scientific knowledge and the variables included in the Logit specification of the ZIP or ZINB.

Table 6. Results of ZIP models

	Poisson											
	Model I		Model II		Model III		Model IV		Model V		Model VI	
	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.
ONE	0.641	0.664	0.874	0.816	0.766	0.684	−0.462	0.623	−0.190	0.645	0.271	0.704
L	1.340*	0.269	−0.015	0.247	1.372*	0.271	0.504*	0.251	0.569*	0.253	0.957*	0.266
D	0.040**	0.022	0.050*	0.025	0.043**	0.023	0.015	0.023	0.019	0.023	0.032	0.022
C	0.186*	0.046	0.349*	0.058	0.173*	0.046	0.377*	0.049	0.323*	0.048	0.214*	0.048
TK	−0.051*	0.006	−0.038*	0.007	−0.052*	0.006	−0.035*	0.006	−0.038*	0.006	−0.047*	0.006
G1	1.072	0.745	1.747	0.855	1.098	0.757	1.380*	0.712	1.236*	0.730	1.016	0.782
G2	0.732	0.678	1.316	0.804	0.740	0.691	1.043	0.636	0.980	0.655	0.807	0.715
G3	1.574*	0.663	2.097*	0.791	1.581*	0.676	1.875*	0.621	1.803*	0.643	1.631*	0.703
G4	0.586	0.678	1.104	0.808	0.601	0.693	0.849	0.640	0.754	0.662	0.589	0.718
S			−0.276*	0.042	−0.019	0.013						
I			−0.056	0.829			0.807*	0.049				
E			0.526	0.413					0.264*	0.018		
O			0.560	3.023							2.761*	0.319
	Logit											
ONE	5.026*	0.615	7.917*	1.561	7.192*	1.084	6.116*	1.058	6.017*	0.960	5.798*	0.824
G1	−0.760	0.941	−0.441	1.061	−0.622	0.958	−0.658	0.973	−0.669	0.959	−0.672	0.954
G2	−1.974*	0.743	−1.650**	0.881	−1.817*	0.744	−1.812*	0.787	−1.812*	0.779	−1.838*	0.755
G3	−4.367*	0.634	−4.181*	0.784	−4.374*	0.637	−4.314*	0.695	−4.343*	0.682	−4.355*	0.649
G4	−1.793*	0.767	−1.600*	0.896	−1.806*	0.782	−1.727*	0.802	−1.770*	0.791	−1.797*	0.777
C	−2.108*	0.597	−2.286*	0.569	−2.178*	0.588	−2.252*	0.581	−2.237*	0.583	−2.139*	0.599
S			−0.471*	0.234	−0.306*	0.123						
I			−1.022	2.984			−1.197**	0.618				
E			0.027	1.493					−0.459*	0.226		
O			10.340	12.749							−6.917**	3.949
−lnL	628.97		589.03		624.10		606.42		610.86		622.91	
V (1)	6.32		6.85		6.42		6.66		6.59		6.40	
LR (2)			79.68									
N obs.	1139		1139		1139		1139		1139		1139	

* Sign. 5%; ** Sign.10%; (1) Vuong statistic; (2) LR = −2(lnL$Model$I−ln$Model$II)

Table 7. Results of ZINB models

	Negative binomial											
	Model I		Model II		Model III		Model IV		Model V		Model VI	
	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.
ONE	−0.298	1.254	−2.647	1.771	0.095	1.620	−1.156	1.037	−0.820	1.037	−0.470	1.217
L	1.966	4.575	1.785	4.071	2.081	4.405	1.423	3.937	1.436	4.033	1.644	4.363
D	0.236	0.225	0.260	0.223	0.257	0.246	0.132	0.210	0.135	0.216	0.175	0.228
C	0.457	0.583	0.605	0.556	0.439	0.550	0.563	0.588	0.492	0.590	0.408	0.612
TK	0.021	0.046	0.044	0.048	0.017	0.049	0.022	0.043	0.017	0.044	0.014	0.047
G1	1.469	2.235	2.732	2.367	1.609	2.313	1.654	2.142	1.516	2.162	1.397	2.225
G2	0.864	1.336	1.895	1.580	0.871	1.385	1.094	1.046	1.049	1.110	0.946	1.283
G3	1.723	1.231	2.694*	1.458	1.766	1.286	1.960*	0.922	1.903*	1.000	1.802	1.184
G4	0.266	1.523	1.293	1.695	0.444	1.607	0.535	1.281	0.466	1.357	0.363	1.509
S			−0.401	0.578	−0.067	0.150						
I			10.977	7.459			0.818	0.698				
E			−4.733	3.937					0.239	0.214		
O			35.495	36.799							2.119	3.195
Alpha	1.473*	0.301	1.517	0.304	1.564*	0.316	1.288*	0.337	1.310*	0.335	1.376*	0.329
ONE	4.467*	0.544	6.658*	1.242	7.259*	0.898	5.650*	0.941	5.596*	0.857	5.331*	0.696
G1	−0.481	0.655	0.254	0.667	−0.287	0.671	−0.423	0.620	−0.445	0.605	−0.404	0.640
G2	−1.850*	0.627	−1.189*	0.618	−1.702*	0.645	−1.699*	0.598	−1.694*	0.598	−1.692*	0.617
G3	−4.201*	0.557	−3.750*	0.543	−4.260*	0.579	−4.170*	0.542	−4.203*	0.541	−4.178*	0.548
G4	−1.812*	0.618	−1.268*	0.633	−1.758*	0.632	−1.746*	0.596	−1.789*	0.586	−1.783*	0.602
C	−2.714*	0.807	−3.234*	0.982	−3.042*	0.854	−2.808*	0.861	−2.823*	0.863	−2.679*	0.817
S			−0.587*	0.100	−0.395*	0.092						
I			2.247	3.900			−1.210*	0.525				
E			−1.538	1.949					−0.484*	0.197		
O			23.258	13.115							−7.474*	3.124
−lnL	468.23		456.85		463.03		463.22		463.74		465.83	
V (1)		3.76		4.21		4.01		3.52		3.61		3.80
LR (2)			*	22.76								
N obs.	1139		1139		1139		1139		1139		1139	

* Sign. 5%; ** Sign.10%.
(1) Vuong statistic
(2) LR = −2(lnL*Model*I−ln*Model*II)

Table 8. Testing of hypotheses: summary of results (1)

Hypotheses	Explanatory variables	POISSON		NEGATIVE BINOMIAL		ZIP				ZINB			
		Relation	Sign (2)	Relation	Sing (2)	Relation		Sing (2)		Relation		Sing (2)	
						P	L	P	L	NB	L	NB	L
H1	S. I. E. O	? (1)	*	? (1)	*		? (1)		*		? (1)		*
H2	S	+		+	*	–	+		*	–	+		*
H3	I	+	*	+	*	+	+	*	**	+	+		*
H4	E	+	*	+	*	+	+	*	*	+	+		*
H5	O	+		+	*	+	+	*	**	+	+		*

(1) : The colinearity prevents the determination of the correct signs when all the variables are introduced simultaneously.
(2) * Sign. 5%; ** Sign.10%.

5. Conclusions

In this study we have outlined some regional trends in the science-technology relationships, for the case of the Spanish regions. We have also identified the factors influencing the utilisation of science by industry, paying special attention to the role played by the regional scientific environment. This topic is especially relevant because, by correctly identifying the elements that encourage (or limit) the transfer of knowledge from the scientific community to industry, the appropriate scientific and technological policy measures can be applied to strengthen the relationships between the scientific community and the industrial sector.

From the descriptive analysis of the data and the application of indicators of concentration, it can be concluded that in Spain there is a strong regional polarisation of science-technology flows. Most of the total science-technology flows are concentrated in only three of the seventeen Spanish regions (Catalonia, Madrid and Navarre, in this order). These regions account for 69.53% of the total NPC citations, and 61% of total scientific citations (ISI). This concentration is greater in citations than in the number of patents (these same three regions account for 62.93% of the total patents), evidence that the concentration of science-technology flows is even more polarised regionally in Spain than technology itself. However, a direct relationship does not exist between technology patented and science technology flows; regions such as Navarre that in a particular sector are unrepresentative of Spain as a whole, are leaders in the application of scientific knowledge to a particular field of technology.

In order to identify the principal factors influencing the application of scientific research by regional industry, a "function of utilisation of scientific knowledge" has been specified and estimated by means of different econometric modellings, in respect of a sample of 1.139 Spanish companies. From this analysis we are able to conclude that the scientific environment is, in general, relevant and determine that the companies located in the regions that invest more in scientific resources and generate more research results, make greater use of science in their patented technology. This demonstrated tendency contributes, in part, to explaining the regional polarisation in science technology relationships to which we have made previous reference. The econometric estimations have also allowed us to refine the conclusions we can draw; for example, the variables found to be statistically related to the scientific citations in patents in most of the econometric specifications are those of the human resources (numbers of researchers in universities, and of graduates from science and technology courses), rather than the variables of scientific research expenditure and results. Further, the ZIP and ZINB specifications that contemplate two regimes, one of access to scientific knowledge and other of intensity in its application, suggest that the nature of the scientific environment is more related to access to scientific research than to intensity in the application of that research.

With the precautionary reservations that must necessarily be accepted when the scientific citations in patent documents are used to reflect science-technology relationships, the results obtained are, therefore, favourable to the hypothesis that larger allocations of regional expenditure and human resources to university research, and the production of more scientific research results and more human resources capable of generating and applying scientific knowledge, contribute to increasing the use of science by regional industry. In consequence, by way of a final reflection, it can be stated that the efforts being made by many regional governments to strengthen university research in the expectation that part of this research will be applied productively by industry, are bearing fruit.

However, this study examining the effect on scientific citations of the amount of resources allocated and the results obtained from university research, has in no case attempted to compare the efficiency or yield from a given amount of resources between different regions. The analysis of this efficiency is more complex in terms of data collection and will be the subject of a future study.

Despite everything, and connected with this last comment, even though higher research expenditure or increased resources are known to generate improved science-technology flows, we should not forget that, if these are not complemented with an integral regional planning of R&D that includes the other elements of the system of innovation and their interrelationships, and an effective coordination with the national and European planning, then the *regional paradox* (regions that are achieving high levels of research excellence present some very poor results in the field of innovation and the technological development of their companies) that is currently found among the different Spanish regions will not be resolved.

References

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

Agrawal A (2001) University-to-industry knowledge transfer: literature review and unanswered questions. International Journal of Management Reviews 3: 285–302

Anselin L, Varga A, Acs ZJ (1997) Local geographic spillovers between university research and high technology innovations. Journal of Urban Economics 42: 422–448

Anselin L, Varga A, Acs ZJ (2000) Geographic and sectoral characteristics of academic knowledge externalities. Papers of Regional Science 79: 435–443

Audretsch DB, Feldman MP (1996) R&D Spillovers and the geography of innovation and Production. American Economic Review 86: 630–640

Audretsch DB, Stephan P (1996) Company-scientist locational links: The case of biotechnology. American Economic Review 86

Bloom N, Van Reenen J (2000) Real options, patents, productivity and market value: evidence from a panel of British firms. The Institute for Fiscal Studies, Working Paper 00/21

Blundell R, Griffith R, Van Reenen J (1999) Market share, market value and innovation in a panel of British Manufacturing Firms. Review of Economic Studies 66: 529–554

Branstetter L (2001) Exploring the link between academic science and industrial innovation: The case of California's research universities. 2001 NBER Conference

Cameron AC, Trivedi PK (1990) Regression based test for overdispersion in the Poisson model. Journal of Econometrics 46: 347–364

Cameron AC, Trivedi PK (1986) Econometric models based on count data: Comparisons and applications of some estimators and tests. Journal of Applied Econometrics 1: 29–54

Cameron AC, Trivedi PK (1998) Regression analysis of count data. Cambridge University Press, New York

Cockburn IM, Henderson RM (1998) Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. The Journal of Industrial Economics 46: 157–182

Cohen WM, Levinthal DA (1989) Innovation and learning: the two faces of R&D. The Economic Journal 99: 569–596
Cohen WM, Levinthal DA (1990) Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35: 128–152
Collins P, Wyatt S (1988) Citations in patents to the basic research literature. Research Policy 17: 65–77
Crepon B, Duguet E (1997) Research and development, competition and innovation. Pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity. Journal of Econometrics 79: 355–378
Eicker F (1967) Limit theorems for regressions with unequal and dependent errors. In: LeCam L, Neyman J (eds) Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press Berkeley, pp. 59–82
Feldman M, Florida R (1994) The geographic sources of innovation: Technological infrastructure and product innovation in the United States. Annals of the Association of American Geographers 84: 210–229
Fischer M, Varga A (2003) Spatial knowledge spillovers and university research: Evidence from austria. Annals of Regional Science 37: 303–322
Fritsch M, Schwirten C (1999) Enterprise-University cooperation and the role of public research institutions in regional innovation systems. Industry and Innovation 6: 69–83
Goe R, Lentneck B, MacPherson A, Phillips D (2000) Toward a contact-based theory of producer service location. Environment and Planning A 32: 131–147
Greene W (1994) Accounting for excess zeros and sample selection in poisson and negative binomial regression models. Working Paper n° EC-94-10, Department of Economics, Stern School of Business, New York University
Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth. Bell Journal of Economics 10: 92–116.
Griliches Z (1981) Market value, R&D and patents. Economic Letters 7: 183–187
Grupp H, Schmoch U (1992) Perception of scientification as measured by referencing between patents and papers. In: Grupp H (ed) Dynamics of Science-Based Innovation. Springer Berlin, Heidelberg New York, pp. 73–128
Hall B, Jaffe A, Trajtenberg M (2001) Market value and patent citations: a first look. NBER Working Paper 7741
Jaffe A (1989) The real effects of academic research. American Economic Review 79: 957–970
Jaffe A, Fogarty M, Banks (1998) Evidence from patents and patent citations on the impact of NASA and other labs on commercial innovations. Journal of Industrial Economics 46 (2)
Jaffe A, Trajtenberg M, Henderson R (1993) Geographic localisation of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108: 577–598
Jaffe A, Trajtenberg M (1996) Flows of knowledge from Universities and Federal Labs: modelling the flow of patent citations over time and across institutional and geographic boundaries. NBER working paper n° 5712
Klevorick AK, Levin R, Nelson R, Winter S (1995) On the sources and significance of inter-industry differences in technological opportunities. Research Policy 24: 342–349
Lambert D (1992) Cero-inflated poisson regression, with an application to the defects in manufacturing. Technometrics 34: 1–14
Lim K (2000) The many faces of a asortive capacity: Spillovers of Copper interconnect technology for semiconductor chips. Mimeo, MIT
Lundvall B, Borras S (1997) The globalization learning economy: Implications for innovation policy. Report for the DG XII Commission
MacPherson A (2002) The contribution of academic-industry interaction to product innovation: The case of New York State's medical device sector. Papers in Regional Science 81: 121–129
Mansfield E (1991) Academic research and innovation. Research Policy 20: 1–12
Mansfield E (1998) Academic research and industrial innovation: an update of empirical findings. Research Policy 26: 773–776
Maurseth PB, Verspagen B (1999) Knowledge spillovers in Europe. A patent citation analysis. Paper presented at the CRENOS conference on "Technological externalities and spatial localisation"
Mcmillan GS, Narin F, Deeds DL (2000) An analysis of the critical role of public science in innovation: the case of biotechnology. Research Policy 29: 1–8

Meyer M (2000a) Does science push technology? Patents citing scientific literature. Research Policy 29: 409–434
Meyer M (2000b) What is special about patent citations? Differences between scientific and patent citations. Scientometrics 49: 93–123
Meyer M (2000c) Patent citations in a novel field of technology. What can they tell about interactions between emerging communities of science and technology? Scientometrics 48: 151–178
Meyer M (2002) Tracing knowledge flows in innovation systems. Scientometrics 54 (2): 193–212
Meyer-Krahmer F, Schmoch U (1998) Science-based technologies: university-industry interactions in four fields. Research Policy 27: 835–851
Mullahy J (1986) Specification and testing of some modified count data models. Journal of Econometrics 33: 341–365
Narin F, Hamilton KS, Olivastro D (1995) Linkage between agency-supported research and patented industrial technology. Research Evaluation 5: 183–187
Narin F, Hamilton KS, Olivastro D (1997) The increasing linkage between US technology and public science. Research Policy 26: 317–330
Narin F, Olivastro D (1992) Status report: linkage between technology and science. Research Policy 21: 237–249
Narin F, Olivastro D (1998) Linkage between patents and papers: an interim EPO/US comparison. Scientometrics 41: 51–59
Noyons ECM, Van Raan AFJ, Grupp H, Schmoch U (1994) Exploring the science and technology interface: inventor-author relations in laser medicine research. Research Policy 23: 443–457
Pavit K (1998) The social shaping of the national science base. Research Policy 27: 793–805
Salter AJ, Martin BR (2001) The economic benefits of publicly funded basic research: a critical review. Research Policy 30: 209–532
Schartinger D, Rammer C, Fischer MM, Fröhlich J (2002) Knowledge interactions between universities and industry in Austria: sectoral patterns and determinants. Research Policy 31: 303–328
Schmoch U (1993) Tracing the knowledge transfer from science to technology as reflected in patent indicators. Scientometrics 26: 193–211
Tijssen RJW (2001) Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. Research Policy 30: 35–54
Tijssen RJW (2002) Science dependence of technologies: evidence from inventions and their inventors. Research Policy 31: 509–526
Tijssen RJW, Buter RK, Van Leeuwen TN (2000) Technological relevance of science: An assessment of citation linkages between patents and research papers. Scientometrics 47: 389–412
Van Vianen BG, Moed HF, Van Raan AFJ (1990) An exploration of the science base of recent technology. Research Policy 19: 61–81
Varga A (1998) University research and regional innovation: A spatial econometric analysis of academic technology transfers. Kluwer Academic Publishers, Dordrecht
Verbeek A, Debackere K, Luwel M, Andres P, Zimmermann E, Deleus F (2002) Linking science to technology: using bibliographic reference in patents to build linkage schemes. Scientometrics 54: 339–420
Verspagen B, Schoenmakers W (2000) The spatial dimension of knowledge spillovers in Europe: Evidence from firm patenting data. AEA Conference, Alicante, Spain
Vuong Q (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57: 307–334
White H (1982) Maximun likelihood estimation of misspecified models. Econometrica 50: 1–25
Zucker L, Darby M, Armstrong J (2000) University science, venture cpital, and the performance of U. S. biotechnology firms. Mimeo, University of California, Los Angeles