Paper to be presented at the DRUID 2011

on

INNOVATION, STRATEGY, and STRUCTURE -
Organizations, Institutions, Systems and Regions
at
Copenhagen Business School, Denmark, June 15-17, 2011

Regional scientific capacities in Europe. Specialization and determining factors

Manuel Acosta
University of Cadiz
Economics
manuel.acosta@uca.es

Daniel Coronado
University of Cadiz
Economics
daniel.coronado@uca.es

Esther Ferrándiz

Economics
esther.ferrandiz@uca.es

M.Dolores León

marilo.leon@uca.es

Abstract

A problem with the max legth was found when introducing the abstract.

Jelcodes:I23,R19

Regional scientific capacities in Europe. Specialization and determining factors

Abstract: Despite the high amount of public funds allocated to academic research and development (R&D), little is known about the effects of such investments at regional level. Additionally, governments must decide scientific priorities and the allocation of funds to regions for which they should have information about the differences in regional scientific capacities. This paper aims shed light on these issues by dividing our research questions into two groups. The first group encompasses some questions related to the regional production of science in Europe. The second group aims to analyse its determining factors, with a special focus on the role of academic R&D expenditures.

The main data set to capture the regional scientific production consists of a regionalization of roughly 1,000,000 papers from 1998 to 2004 obtained from the Thomson ISI (Information Sciences Institute) database. The methodology involves a descriptive analysis and several econometric models. The empirical model is based on a regional knowledge production function estimated with an unbalanced sample of European regions for the period 1998 to 2004. Usual panel techniques and several sensibility analyses are applied to obtain robust results.

1. Introduction

While there is some evidence of the short-term usefulness of incentives at country level to promote scientific research, regional information seems to be largely missing in the literature. In this paper we try to fill this gap providing some insight on the distribution and academic scientific specialization across European regions; additionally, we particularly focus on the role financial funds on the development of university-based research at regional level. More specifically our research questions are as follows:

1. How is the production of scientific research distributed across European regions? Which regions lead in scientific production by discipline? What are the regional scientific specialization patterns across European regions?.

2. What are the effects of academic R&D funding in promoting the production of scientific research at regional scale? Is there any difference according to the level of regional development?

There are several major arguments that prompt this research. First, the identification of academic scientific capacities is an important issue for regional, national and supranational governments that must decide scientific priorities and the allocation of funds. Second, knowing the underlying mechanisms to promote scientific research is useful for their implications in the development of industrial innovations (for a review see Cohen et al., 1998). Our interest in regions as units of analysis stems from both the growing importance of knowledge produced in universities for regional innovation systems and the need to consider the differences in scientific performance and specialisation profiles across regions, given that -at least in Europe- policy competences and institutions are partly bound to regions (Cooke et al., 2000).

We address our research question by using detailed data on academic published papers by regions obtained from the ISI Web of Knowledge database. The other important variable is Higher Education R&D expenditures (HERD). It is worth remarking that this study is an exploratory research conditioned by data limitations. First, despite the ISI database has been frequently used in several bibliometric analyses, but it has some well known drawbacks. Second, from the point of view of measuring basic research, academic R&D statistics for European regions are subject to considerable lack of accuracy; the available data is aggregated and does not consider all the resources received by universities. The statistics present also a problem of missing observations for many European regions.

The paper is organized as follows. Section 2 summarizes the empirical literature relevant to this paper. Section 3 describes the data and provides an overview on the patterns of university scientific production at regional level across Europe 15. Section 4 presents a regional version of a knowledge production function. The estimations of several models explaining the effects of HERD on the production of scientific knowledge are provided in Section 5. We briefly summarize the conclusions in the final Section.

2. Empirical background

Despite the theoretical and empirical relevance of several economic aspects involving research activities in universities (see the surveys by Dasgupta and David, 1994; Stephan, 1996), the empirical analysis concerning the production of science for universities is very scarce; only a few papers have addressed this issue from an economic view.[1] In this brief review, we summarize previous results relevant to our research; particularly our focus of attention is twofold: i) what is the role of academic R&D expenditures in promoting academic scientific production?; ii) what other factors may determine the production of scientific research?.

A couple of papers by Adams and Griliches (1996, 1998) were one of the early attempts to measure the relation between inputs (R&D expenses) and outputs (scientific publications and citations) from an economic view. Their point of departure was the evidence of a discrepancy between the growth of R&D expenses (5.5% per year in real terms) and the total number of scientific articles (1% per year) in the United States during 1981-1991. Several regressions using different lags for R&D provided average elasticity of 0.6 for papers and 0.7 for citations at the university and field level, suggesting the possibility of diminishing returns to scale, but not at the aggregate level where constant returns was the result. However, the results were possibly biased because, as the author remarked, spillover effects among universities and fields were not taken into account and the difference in elasticities with more aggregated models was possible there. Therefore, serious data limitations and difficulties hindered the authors from drawing firm conclusions. In subsequent research, Adams et al. (2005) studied the size of scientific teams and institutional collaboration with data about 2.4 million scientific papers written in 110 top U.S. research universities that have at least one author form this set of leading US universities. Their analysis was carried out over the period 1981-1999. The source of their data was the Institute for Scientific Information (ISI). Although Adams et al. (2005) addressed several questions related to collaboration and team size, a part of the econometric analysis includes the determinants of research outputs, where they found positive and highly significant coefficients of the logarithm of the lagged stock of R&D (with elasticity around 0.45 for papers, and 0.55 for citations, suggesting diminishing returns to the stock of R&D applied at the university-field level. Payne and Siow (2003) estimated the effects of federal research funding on research outcomes in the U.S. at university level. They used data from the Institute for Scientific Information (ISI) about articles published and the counts of citations received. Their analysis covered 57 universities and 1,017 observations, representing about 18 years of data for each university. Using scientific articles as the dependent variable, all their estimations for federal research funding were significant and showed diminishing returns. They also used citations per article but obtained a negative and very small effect. Consequently, the authors concluded that increasing federal research funding results in more, but not necessarily higher quality research output. Following a similar methodology to Adams and Griliches (1998), Crespi and Geuna (2008) examined the determinants of scientific production of 14 OECD countries in the period 1981-2002.

---

[1] A stream of literature focused on the individual productivity of researchers has sometimes considered R&D funding as an “environmental attribute”, along with other personal characteristics (researcher sex, age, education, etc.) and institutional attributes of the institutions for which those researchers work (see for a review, Carayol and Matt, 2006). However, this growing literature is far from the view followed in our paper.

The outputs were taken from the Thomson ISI national science indicators database on published papers and citations. This research differs form Griliches and Adams (1998) in several points (different structure of data, context, etc.), but mainly in considering spillover effects of HERD in the original knowledge production function. Assuming the difficulties to obtain robust results for elasticities of the outputs (number of papers and citations), given the poor quality of the data and modelling problems, their models suggested decreasing returns to the domestic component of R&D. The analysis of international spillovers indicated evidence of a significant impact from the weighted investment in HERD in other countries. They also concluded that there are no significant effects from past R&D expenditure on publication output after six years.

3. Data

This Section is divided into three parts. The first part summarizes the data collection process and the statistical sources. The second part describes the spatial distribution of academic research and the specialization patterns from 1998 to 2004 across EU-15 regions. The third part provides a descriptive overview on the relationships between inputs (academic R&D expenditures) and outputs (scientific production) at the regional level.

3.1 Data collection process

The empirical data used in this study consists of a set of university research articles published in scientific journals indexed by the Science Citation Index Expanded (SCI). As is well known, the SCI is a bibliographical database produced by the Information Sciences Institute (ISI), which is in turn a part of Thomson Reuters' Web of Science. The main advantage of ISI citation indexes is that they provide a complete list of all authors and their affiliations. There are also some known limitations of this database. For example, it does not include all journals, and the ISI journal list is strongly biased towards journals published in English.

The procedure to account for university scientific papers at NUTS II level in the EU-15 followed three steps:

(i) Data on academic publications containing at least one author affiliated with a university from an EU-15 country for 1998–2004 were retrieved from the SCI. It is worth emphasizing that the lack of normalization in the way in which academic institutions are named hinders the finding of academic publications. For this reason, we included several search terms to help identify Higher Education institutions in both English and other languages (fachhochschule, yliopisto, ecole, institut nacional polytehcnique, politécnico, scuola, hogskola, etc.). This search resulted in 994,938 publications.

(ii) Regionalization at the NUTS II level of aggregation of the academic publications obtained in Step 1 (213 regions[2]). We first identified the NUTS II associated with each university using the list provided by the members of

---

[2] Number of regions in the EU-15 according to Regulation (EC) No. 1059/2003 of the European Parliament and the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS) (excluding extra-regio).

the European Indicators, Cyberspace and the Science-Technology-Economy System (EICSTES). For those universities not included in the EICSTES list, we searched for the address on each university's website. We then applied the full-count method involving the crediting of each co-author with one publication. Following the full-counting procedure, the total number of academic papers distributed among regions during 1998–2004 was 1,206,644.

(iii) Classification of academic publications by scientific field. In a first stage, we classified the 7,155 journals in our sample according to the categories defined by the Journal Citations Report (JCR) database for the ISI. However, this classification is generally too specific, because it includes more than two hundred individual categories. For this reason, in a second stage we grouped the ISI categories into 12 broad scientific disciplines using the Third European Report on S&T indicators[3]. In this classification, each ISI category is assigned to only one scientific discipline, but each journal is assigned to several categories by the ISI. If a journal was assigned to more than one scientific discipline, we applied the full-count method so that we considered only a single full publication for each discipline.

Following this procedure, the top five scientific fields in terms of publications accounted for 67.94% of the total number of scientific publications in EU-15. These scientific fields are Clinical Medicine (17.33% of total number of publications), Physics and Astronomy (14.66%), Chemistry (12.11%), Biomedical Sciences (11.98%) and Basic Life (11.86%).

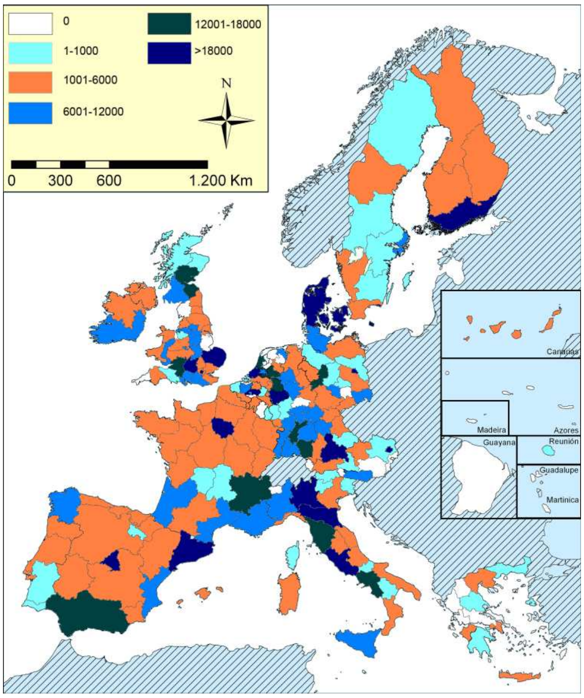
3.2 Scientific capacities and specialization patterns of European regions

The spatial distribution of publications is mapped in Figure 1. As shown, of the 213 regions in total, 24 do not have any scientific publications, 34 have between one and 1,000 publications, 73 have from 1,000 to 6,000 publications, 43 have from 6,000 to 12,000 publications, 16 have from 12,000 to 18,000 publications and 23 have more than 18,000 publications. Figure 1 also reveals that German and Italian regions appear the most likely to publish in scientific journals.

---

[3] The classification was established by the Centre for Science and Technology Studies (CWTS) at Leiden University (see Tijssen and van Leeuwen 2003). For categories not included in the CWTS 2003 classification, we used an updated (but unpublished) classification kindly provided by the CWTS.

Figure 1. Spatial distribution of academic scientific papers by NUTS 2 in Europe 15 (1998-2004)



Source: Own elaboration based on ISI data

The indexes in Table 1 reveal that the production of scientific knowledge is highly concentrated in a few regions. As shown in Table 1, the Gini coefficient takes a value of 0.61 for the initial year (1998) and 0.59 for the latest year (2004) in the sample. Moreover, as shown, the trend in the Gini coefficients is slightly downward over the period 1998 to 2004. The remaining concentration indexes in Table 1 lead to the same conclusion. For example, the value of the C5 index takes a value of about 13, suggesting that just five regions account for 13% of papers. Similarly, the value of the C10 index is 22, indicating that 10 regions provide almost 22% of publications.

Table 1. Descriptive statistics and regional concentration indexes of academic scientific publications

	1998	1999	2000	2001	2002	2003	2004	98–04
N	157,446	164,492	166,66	170,603	174,266	179,770	193,398	1,206,644
Mean	739.19	772.27	782.49	800.96	818.16	844.00	907.98	5,664.99
Max.	5,794	5,950	5,887	6,162	6,186	6,401	6,701	43,081
Min.	0	0	0	0	0	0	0	0
Std. Dev	937.41	972.52	976.27	995.60	1,013.52	1,046.23	1,100.44	7,024.09
C. Var(1)	1.26	1.25	1.24	1.24	1.23	1.23	1.21	1.23
Coeff. Gini(2)	0.61	0.60	0.60	0.60	0.59	0.59	0.58	0.59
C5(3)	13.38	13.38	13.16	13.31	13.12	13.32	12.91	13.18
C10(4)	23.04	22.88	22.82	22.65	22.41	22.64	22.13	22.61
C25(5)	44.86	44.69	44.13	44.05	44.10	44.06	43.37	44.02

(1) Coefficient of variation = Std Dev. ÷ Mean; (2) The Gini coefficient ranges between 0 and 1; the larger the value the higher the level of regional concentration in publications or collaborations. (3)(4)(5) Concentration indexes of publications for the top 5, 10 and 25 regions with the largest number of scientific papers, respectively.

Table 2 provides details on the 10 regions with the highest rate of publications. Of these 10 leading regions, United Kingdom accounts for 3 regions and German for 2 regions. The top ten regions in terms of publications accounted for 22.61% of total number of publications, which confirms that scientific knowledge production is highly concentrated in a few regions. Note that after data normalization only UK regions remain in the top ten ranking.

Table 2. Regions with the highest number of academic publications and academic publications per capita (Annual average 1998-2004)

	Nº Papers	%	Cum. (%)	Scientific publications/ Population (in thousands)	
Île de France (FR10)	6,154	3.57	3.57	Inner London (UKI1)	3.01
Inner London (UKI1)	5,715	3.32	6.89	Prov. Vlaams-Brabant (BE24)	2.62
Denmark (DK00)	3,915	2.27	9.16	Berkshire, Buck. and Oxf. (UKJ1)	2.28
Oberbayern (DE21)	3,543	2.06	11.21	East Anglia (UKH1)	2.23
Lombardia (ITC4)	3,395	1.97	13.18	Wien (AT13)	1.85
Etelä-Suomi (FI18)	3,384	1.96	15.14	Groningen (NL11)	1.81
Berlin (DE30)	3,344	1.94	17.08	Eastern Scotland (UKM2)	1.76
East Anglia (UKH1)	3,296	1.91	19	Kärnten (AT21)	1.73
Berkshire, Buck. and Oxf. (UKJ1)	3,233	1.88	20.87	Gießen (DE72)	1.63
Cataluña (ES51)	2,989	1.73	22.61	Utrecht (NL31)	1.61
Others	133,409	77.39	100		
Annual average.	172,378	100			

Table 3 lists the regions with the highest level of scientific production by discipline according to their share of total publications in each scientific field. Again, for every discipline the greatest numbers of publications are concentrated in just a few regions. The level of concentration of the 25 regions with more publications ranges from 40% to 50%. Note that Île de France appears as a top-five region in nine of the 12 scientific disciplines. It is also remarkable that Inner London (UK) appears as a top-five region across all of the scientific disciplines.

Table 3. Regions with the highest rate of publications by scientific discipline and indexes of concentration (1998-2004)

Scientific field	Regions	No. of regional papers ÷ No. of papers (%)	Concentration indexes
Agricultural and Food Sciences	Denmark (DK00) Gelderland (NL22) Oberbayern (DE21) Etelä-Suomi (FI18) Inner London (UKI1)	4.29 3.99 2.85 2.62 2.57	C5 (16.32) C10 (27.65) C25 (52.24)
Basic Life	Inner London (UKI1) Île de France (FR10) Denmark (DK00) Etelä-Suomi (FI18) East Anglia (UKH1)	3.65 3.15 2.99 2.2 2.19	C5 (14.18) C10 (24.40) C25 (46.68)
Biological Sciences	Denmark (DK00) Île de France (FR10) Inner London (UKI1) Gelderland (NL22) Berk., Buc. and Oxf. (UKJ1)	3.59 2.8 2.74 2.39 2.26	C5 (15.94) C10 (25.52) C25 (45.29)
Biomedical sciences	Inner London (UKI1) Lombardia (ITC4) Berlin (DE30) Denmark (DK00) Oberbayern (DE21)	4.33 2.66 2.53 2.51 2.47	C5 (14.5) C10 (25.4) C25 (49.09)
Chemistry	Île de France (FR10) Cataluña (ES51 Inner London (UKI1) Emilia-Romagna (ITD5) Oberbayern (DE21)	3.93 2.31 2 1.97 1.96	C5 (12.17) C10 (20.89) C25 (40.95)
Clinical medicine	Inner London (UKI1) Lombardia (ITC4) Oberbayern (DE21) Etelä-Suomi (FI18) Zuid-Holland (NL33)	4.1 3.39 2.91 2.69 2.66	C5 (15.75) C10 (27.77) C25 (51.60)
Computer Sciences	Île de France (FR10) Inner London (UKI1) Denmark (DK00) Cataluña (ES51 Andalucía (ES61)	3.64 3.27 2.13 2.03 1.92	C5 (12.99) C10 (21.98) C25 (42.56)
Earth Sciences	Île de France (FR10) Denmark (DK00) Inner London (UKI1) East Anglia (UKH1) Berk., Buc. and Oxf. (UKJ1)	3.73 2.95 2.93 2.8 2.78	C5 (15.19) C10 (24.65) C25 (43.58)
Engineering sciences	Île de France (FR10) Inner London (UKI1) Zuid-Holland (NL33) East Anglia (UKH1) Rhône-Alpes (FR71)	3.58 3.44 2.17 1.95 1.81	C5 (12.95) C10 (21.34) C25 (41.1)
Mathematics and Statistics	Île de France (FR10) Inner London (UKI1) Andalucía (ES61) Lazio (ITE4) Comunidad de Madrid (ES30)	8.1 2.43 2.35 2.11 2.1	C5 (17.09) C10 (25.84) C25 (44.79)

Physics and Astronomy	Île de France (FR10) East Anglia (UKH1) Inner London (UKI1) Rhône-Alpes (FR71) Lazio (ITE4)	5.47 2.9 2.82 2.15 1.98	C5 (15.32) C10 (24.80) C25 (45.31)
Multidisciplinary	Inner London (UKI1) East Anglia (UKH1) Berk., Buc. and Oxf. (UKJ1) Île de France (FR10) Denmark (DK00)	5.41 5.39 5.14 4.66 2.45	C5 (23.05) C10 (32.66) C25 (51.35)
(*) Concentration indexes of publications for the 5, 10 and 25 regions with the largest number of scientific papers.			

Table 4 shows the specialization index of the 20 regions with the highest number of papers in twelve scientific fields. The scientific specialization index is calculated in a similar way to the Revealed Technological Advantage (Soete and Wyatt, 1983) as shown in equation 1.

$$RSA=\frac{Pij/\sum_{j=1}^{12}Pij}{\sum_{i=1}^{12}Pij/\sum_{i=1}^{213}\sum_{j=1}^{12}Pij} \qquad \text{Equation 1.}$$

where *Pij* is the publication counts of region *i* in discipline *j*; *Pj* is the publication counts of region *i* in all disciplines; *Pi* is the publication counts of all regions in discipline *j*; and *P* is the publication counts of all regions in all disciplines. An index greater than one, suggests a relative scientific strength of the region in that specific discipline. For instance, Cataluña (ES51) and Emilia-Romagna (ITD5) are the two regions with the highest RSA in Chemistry; Zuid-Holland (NL33) and Prov. Vlaams-Brabant (BE24) have the highest RSA in Engineering sciences.

Table 4. Scientific specialization of the top 20 research regions[4]

	1	2	3	4	5	6	7	8	9	10	11	12	H (*)
Île de France (FR10)	0.29	0.90	0.80	0.68	1.13	0.52	1.05	1.07	1.03	2.33	1.34	1.57	0.13
Inner London (UKI1)	0.78	1.11	0.83	1.31	0.61	1.25	0.99	0.89	1.05	0.74	1.64	0.85	0.12
DanmarkD (DK00)	1.84	1.28	1.54	1.08	0.69	0.97	0.91	1.26	0.75	0.73	1.05	0.76	0.15
Oberbayern (DE21)	1.42	1.04	0.71	1.23	0.98	1.45	0.63	0.60	0.54	0.51	1.02	0.92	0.17
Lombardia (ITC4)	0.75	1.03	0.56	1.36	0.73	1.73	0.83	0.56	0.64	0.77	0.52	0.79	0.14
Etelä-Suomi(FI18)	1.33	1.11	1.07	1.14	0.71	1.36	0.92	1.05	0.86	0.52	0.69	0.76	0.14
Berlin (DE30)	1.06	1.04	0.82	1.32	0.86	1.38	0.61	0.55	0.53	0.89	0.71	1.03	0.15
East Anglia (UKH1)	0.56	1.19	1.14	0.79	0.79	0.53	0.71	1.52	1.06	0.70	2.94	1.58	0.11

---

[4] Regions with the highest number of publications

Berk., Buck.and Oxf. (UKJ1)	1.11	1.19	1.24	1.03	0.91	0.70	0.86	1.53	0.75	0.82	2.83	1.07	0.11
Cataluña (ES51)	1.25	1.10	1.11	0.91	1.30	0.82	1.14	1.06	0.87	1.15	0.65	0.88	0.11
Zuid-Holland (NL33)	0.24	1.08	0.64	1.24	0.66	1.56	1.12	0.59	1.27	0.69	0.98	0.72	0.10
Köln (DEA2)	0.58	0.97	0.87	1.28	0.78	1.45	0.68	0.66	0.51	0.80	0.77	1.19	0.13
Lazio (ITE4)	0.40	1.05	0.70	1.08	0.69	1.14	1.14	0.71	1.07	1.28	0.73	1.21	0.12
Wien (AT13)	1.51	0.93	0.98	1.16	0.76	1.44	0.97	0.78	0.69	0.78	0.64	0.86	0.13
Emilia-Romagna (ITD5)	0.92	0.95	0.71	1.12	1.26	1.13	0.91	0.71	0.87	0.74	0.50	1.02	0.16
C. de Madrid (ES30)	1.32	1.07	1.05	1.03	1.07	0.62	1.16	0.71	1.00	1.38	0.78	1.19	0.13
Noord-Holland (NL32)	0.38	1.01	0.88	1.39	0.60	1.76	0.97	1.13	0.40	0.61	1.05	0.78	0.1
Prov. Vlaams-Brabant (BE24)	1.30	1.15	0.88	1.09	0.95	1.02	1.03	0.76	1.10	0.91	0.72	0.85	0.13
Karlsruhe (DE12)	0.14	0.81	0.37	1.09	0.96	1.52	1.02	0.89	0.85	0.99	0.88	1.09	0.16
Toscana (ITE1)	0.60	0.92	1.03	1.08	1.02	1.23	1.12	0.96	0.82	1.08	0.53	0.93	0.15

1 Agricultural and Food Sciences; 2 Basic Life; 3 Biological Sciences; 4 Biomedical sciences; 5 Chemistry; 6 Clinical medicine; 7 Computer Sciences; 8 Earth Sciences; 9 Engineering sciences; 10 Mathematics and Statistics; 11 Physics and Astronomy; 12 Multidisciplinary.

(*) Herfindahl Concentration Index of regional publications by scientific disciplines (the index takes value 1 when all the scientific papers published by a region refers to just one discipline, and 1/12 when the regional distribution of papers among scientific fields is the same across all disciplines).

3.2. Scientific production, HERD and level of economic development

Table 5 details the number of academic papers, separating regions with different levels of economic development and HERD. Several facts emerge from this table:

- The distinction between regions according to the level of economic development (GDP per capita) shows that less-developed NUTS regions generated 13.3% of all EU-15 academic papers in 1998. This percentage increased to 15.7% in 2004. On average, an Objective 1 region produced 339 papers in 1998, while a developed NUT region generated 904 papers in the same year. Therefore, the number of academic papers in a less-developed region was 37% of those generated in a developed NUTS region. This figure increased to 45% in 2004.
- By considering that several regions may be included in the group of developed regions, but having a low level of scientific capacity (e.g. regions with a strong tourism sector), the right-hand panel of Table 5 divides regions according to the level of HERD per capita. These data show that regions with less than 75% of the EU-15 average HERD per capita (42% of all regions in the sample) contributed to 12% of all publications in 1998, increasing to 13.4% in 2004. On

average, a region in this group produced 79% fewer papers than a region in the group with HERD per capita higher than 75% of the EU-15 average.

Table 6. Regional production of academic papers by type of NUTs region [(*)]

Groups of regions according to their level of development					Groups of regions according their level of HERD				
		1998	2004	1998–2004 % increase			1998	2004	1998–2004 % increase
Regions below 75% of the EU-15 average GDPpc	No. Papers	20,996	30,463	45.09	Regions below 75% of the EU-15 average HERDpc	No. Papers	12,064	15,855	31.42
	Mean	338.65	491.34			Mean	236.55	317.10	
	Std. Dev	434.08	589.93			Std. Dev	319.20	430.61	
Regions above 75% of the EU-15 average GDPpc	No. Papers	136,450	162,935	19.41	Regions above 75% of the EU-15 average HERDpc	No. Papers	86,905	102,274	17.68
	Mean	903.64	1,079.04			Mean	1,259.49	1,461.06	
	Std Dev	1,035.29	1,211.95			Std Dev.	1,054.00	1,214.21	

(*) The group of less-developed regions comprises 62 NUTS regions, while the number of NUTs regions with more than 75% of the EU-15 average GDP per capita is 151 (213 in total). Because of the lack of data, the number of regions with less and more than 75% of EU-15 average university R&D per capita falls to 120, with 51 NUTS regions in the first group and 69 in the second group.

This analysis shows an unbalanced picture of the generation of academic papers since the average capacity for publication of a less developed region is about 45% of the capacity of a developed region in the core group. The disparities are rather stronger when we consider a classification of regions based on Higher Education R&D expenditures.

4. Model and variables

In order to evaluate the impact of R&D expenditures on science production, we estimate a regional version of the knowledge production function suggested by Adams and Griliches (1996) in terms of inputs and outputs. The inputs are academic R&D funds; the outputs are research publications. The empirical panel model takes the form:

$$SP_{it} = \beta RD(r)_{it} + \alpha_i + \eta_t + u_{it}$$

where the dependent variable SP_{it} is the scientific knowledge production proxied by the logarithm of papers for the region i in the year t. The explanatory variables are as follows:

- RD(r) is the logarithm of a distributed lag function of past R&D expenditures.

- Lambda represents regional specific effects. This variables is included because research activity is affected by several others contextual elements such as cultural practices or particular regional innovation policies.
- Nu represents time effects;
- u represents all other unaccounted forces determining this particular measure of output.

Note that this framework is different to those put forward in previous research in some points. First, we focus on a regional perspective, whilst Adams and Griliches (1996, 1998) analysed a country (USA) and Crespi and Geuna (2008) studied a sample of OECD countries. Second, we also introduce regional demand-side effects as a factor that may encourage or hinder the university scientific production. Our primary interest focuses on the parameters beta and lambda. The first would measure the returns to the scale of the region research effort level, if everything else were correctly specified in this equation, while the second would indicate the changing level of efficiency to convert resources into results.

Adams and Griliches (1998) argued several measurement problems that arise when academic R&D expenditures are included as an explanatory factor. Next, we highlight two relevant problems for this paper:

a) Research process takes times, so it is needed to put forward a lag structure accordingly. The usual proposals in the literature consist of using a three or a five-year lag and applied typical U-shaped function to R&D expenditures (lineal in Adams and Griliches, 1996 and geometrical in Crespi and Geuna, 2008). We tried both in our models.

b) R&D expenditures as input measurement can contain errors. On the one hand, measuring academic R&D may underestimate the total value of R&D resources if there are other financial sources; on the other hand, R&D expenditures overstates the total value of resources devoted to research because some of the R&D is assigned to the production of other outputs (for instance, university patents). We assume this imperfection of R&D expenditure because we do not have a better input indicator (and besides the error can be of both signs).

5. Results

To study the effects of academic R&D funds on regional scientific, we estimated several fixed effects regressions including HERD expenditures with different time lags. Models t-1, t-2, t-3, t-4, t-5 include HERD data lagged 1, 2, 3, 4, 5 years respectively. Models named weight 3t and weight 5t includes the inverted V-lag for R&D in the previous three and five years, respectively.

5.1 Effect of Higher Education R&D expenditures on scientific production

As shown in the Table 7, HERD is only significant for the two-year lag model and for the three-year weighted model. The higher significance of HERD variable in the 2-year lag model than for the 3-year weighted model suggests that, for the regions in the sample, the 2-year lag model is the best representation of the effects of academic R&D expenditures on publications. This result suggests that HERD take two years in promoting scientific production, and then this positive effect disappears. Regional effects and time effects are significant for all the models, which confirm the presence of regional differences in the scientific capability to transform the input in outputs and the existence of a temporal effect, since scientific publications tend to increase over time. The Hausman tests show that the fixed effect model is to be preferred in all the estimated models. The number of observations and regions vary across the models due to the lack of data for HERD in some years.

Table 7. Panel data estimates (Fixed effect) of HERD on the number of scientific papers published by regions (1998-2004)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	t-1	t-2	t-3	t-4	t-5	Weight 3t	Weight 5t
Constant	5.408*** (0.561)	5.954*** (0.086)	5.928*** (0.119)	6.092*** (0.077)	6.020*** (0.102)	5.966*** (0.106)	6.082*** (0.097)
HERD	0.148 (0.127)	0.046** (0.021)	0.044 (0.031)	0.005 (0.014)	-0.011 (0.026)	0.048* (0.025)	-0.001 (0.025)
R^2	0.256	0.280	0.236	0.224	0.245	0.279	0.302
F (reg. effects)	115.56***	227,43***	149.85***	163.64***	192.76***	304.87***	263.21***
F (year effects)	12.14***	17.49***	16.00***	17.01***	19.06***	8.66***	8.68***
Hausman Test	187.73***	242.46***	271.64***	230.67	173.81***	196.22***	162.38***
No. obs.	771	725	762	720	638	549	425
No. regions	174	163	161	158	123	112	104

*p<0.10, **p<0.05, ***p<0.01

Notes: Dependent variable is the number of papers in log. Robust standard errors are in parentheses. All regressions include time dummies. t-1, t-2, ... are the lagged HERD variables for one, two… years. Weight 3t: Is the inverted V-lag with weights 0.25, 0.5, and 0.25, respectively, for HERD lagged one, two, and three years. Weight 5t: Is the inverted V-lag with weights 0.111, 0.222, 0.333, 0.222, 0.111 respectively, for HERD lagged from one to five years. R^2 is the within R^2 for fixed effects. Robust clusters standard errors.

5.2 Centre-periphery models of scientific production

In order to test the centre-periphery hypothesis, Table 8 details the number of academic papers, separating objective 1 regions, those with a GPD below the 75% percent of the Community average, from the rest. Several facts emerge from this table. Higher Education R&D expenditures in more developed regions produced results within the two and three subsequent years. This lag is shorter than for less developed regions, in which HERD take five years to positively affect the amount of scientific publications. However, the impact of such investment is greater for objective 1 regions than for more developed regions. This could be explained because the second group of regions is likely to have a ready-to-use knowledge infrastructure, which allows them to convert faster HERD into new scientific knowledge, but also to be less dependent on new funds than objective 1 regions. Conversely, objective 1 regions are likely to lag in knowledge pool and innovative capacity, which takes time to be developed. This has two main consequences: (1) it takes more time to convert HERD into new knowledge in these regions and (2) they are more dependent on research funding, which could also explain the greater impact of HERD on scientific production in these regions. For both types of regions and for all the estimated models, regional and time effects are significant.

Table 8. Panel data estimates (Fixed effect) of HERD on the number of scientific papers published by type of region (1998-2004)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Objective 1 regions		t-1	t-2	t-3	t-4	t-5	Weight 3t	Weight 5t
	Constant	4.889*** (0.592)	5.414*** (0.227)	5.058*** (0.634)	5.545*** (0.321)	4.484*** (0.301)	5.120*** (0.591)	3.879*** (0.731)
	Coeff. HERD	0.155 (0.154)	0.078 (0.057)	0.142 (0.175)	0.045 (0.081)	0.325*** (0.078)	0.182 (0.145)	0.518*** (0.178)
	R^2	0.513	0.635	0.539	0.643	0.736	0.662	0.883
	F (reg. effects)	120.16 ***	289.72 ***	128.07 ***	284.80 ***	376.88 ***	381.95 ***	888.45 ***
	F (year effects)	13.13***	19.91***	12.16***	20.04***	8.78***	2.84*	4.25***
	Hausman Test	67.41***	50.17***	49.28***	53.26***	36.06***	60.16***	20.00***
	No. obs.	199	173	184	170	157	121	91
	No. regions	45	41	41	40	34	28	22
Non objective 1 regions		(8)	(9)	(10)	(11)	(12)	(13)	(14)
		t-1	t-2	t-3	t-4	t-5	Weight 3t	Weight 5t
	Constant	6.057*** (0.515)	6.034*** (0.091)	6.056*** (0.059)	6.126*** (0.082)	6.199*** (0.112)	5.969*** (0.109)	5.946*** (0.104)
	Coeff. HERD	0.039 (0.112)	0.052** (0.022)	0.051*** (0.015)	0.022 (0.021)	-0.009 (0.029)	0.060** (0.026)	0.037 (0.029)
	R^2	0.152	0.186	0.539	0.133	0.136	0.197	0.203
	F (reg. effects)	123.20 ***	229.62 ***	189.63 ***	150.57 ***	185.12 ***	308.88 ***	254.38 ***
	F (year effects)	4.98***	6.46***	4.02***	5.83***	7.04***	4.53***	2.09*
	Hausman Test	122.60 ***	196.85 ***	229.97 ***	239.76 ***	161.27 ***	150.13 ***	122.95 ***
	No. obs.	566	546	572	544	476	422	329
	No. regions	129	122	120	117	89	84	81

$^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$

Notes: Dependent variable is the number of papers in log. Robust standard errors are in parentheses. All regressions include time dummies. t-1, t-2, ... are the lagged HERD variables for one, two… years. Weight 3t: Is the inverted V-lag with weights 0.25, 0.5, and 0.25, respectively, for R&D lagged one, two, and three years. Weight 5t: Is the inverted V-lag with weights 0.111, 0.222, 0.333, 0.222, 0.111 respectively, for R&D lagged from one to five years. R2 is the within R2 for fixed effects. Robust clusters standard errors.

5.3 Accounting for spillovers

6. Conclusions

Authors' note: This is a paper in progress. Future versions will include spillover effects. Additionally, other explanatory variables may be added to the empirical model.

References

Adams, J.D., Griliches, Z., (1996). Research Productivity in a System of Universities. National Bureau of Economic Research, Inc.

Adams, J.D., Griliches, Z.,(1998). Research Productivity in a System of Universities. Annales d'Economie et de Statistique.

Carayol, N., Matt, M. Individual and collective determinants of academic scientists productivity. Information Economics and Policy 18: 55- 72, 2006

Cohen, W.M., Florida, R., Randazzese, L., Walsh, J.P., (1998). Industry and the Academy: Unease partners in the cause of technological advance in: Noll, R.G. (Ed.), Challenges to Research Universities. Brookings Institute Press, Washington, DC.

Cooke, P., Boekholt, P., Tödtling, F., (2000). The Governance of Innovation in Europe: Regional Perspectives on Global Competitiveness. Pinter, London.

Crespi, G.A., Geuna, A.,(2008). An empirical study of scientific production: A cross country analysis, 1981-2002. Research Policy 37, 565-579.

Dasgupta, P., David, P.,(1994). Towards a new economics of science. Research Policy 23, 487-521.

Payne, A., Siow, A.,(2003). Does Federal Research Funding Increase University Research Output? Advances in Economic Analysis & Policy 3, 1018.

Soete, L., Wyatt, S.,(1983). The use of foreign patenting as an internationally comparable science and technology output indicator. Scientometrics 5, 31-54.

Stephan, P.,(1996). The economics of science. Journal of Economic Literature 34, 1199-1235.

Tijssen, R.J.W, van Leeuwen, T.N., (2003). Bibliometric analyses of world science, extended technical annex to chapter 5 of the third European report on science & technology indicators.