SEWPS
SPRU Electronic Working Paper Series

Paper No. 133

Modelling and Measuring Scientific Production: Results for a Panel of OECD Countries

Gustavo Crespi and Aldo Geuna
(SPRU)

April 2005

SPRU Science and Technology Policy Research

The Freeman Centre, University of Sussex,
Falmer, Brighton BN1 9QE, UK
Tel: +44 (0) 1273 873592
E-mail: g.a.crespi@sussex.ac.uk
http://www.sussex.ac.uk/spru/

Modelling and Measuring Scientific Production: Results for a Panel of OECD Countries

Gustavo Crespi[+] and Aldo Geuna*

SPRU, University of Sussex

10/04/2005

Acknowledgements
The authors are grateful to Ed Steinmueller, Stefano Brusoni, Paul David, Pierre Mohnen and Fabio Montobbio for their comments and suggestions. Earlier versions of this paper were presented at the Schumpeter Society Conference 2004, Milan; DRUID Conference 2004, Copenhagen, the Colloquium on Measuring the Impacts of Science 2004, Montreal. The comments and suggestions of participants at these meetings are much appreciated. The authors would also like to thank Evidence Ltd for supplying some of the data used in the econometric analysis. This paper is derived from a report commissioned by the UK Office of Science and Technology, Department of Trade and Industry. All mistakes and omissions, along with the views expressed, remain the sole responsibility of the authors.

[+]: SPRU-University of Sussex, Freeman Centre, Brighton BN1 9QE, UK; Tel. +44 1273 873592; Fax. +44 1273 685 865 e-mail g.a.crespi@sussex.ac.uk
* *Corresponding author*: SPRU-University of Sussex, Freeman Centre, Brighton BN1 9QE, UK; Tel. +44 1273 877139; Fax. +44 1273 685 865 e-mail a.geuna@sussex.ac.uk

Abstract

This paper presents results from employing an econometric approach to examine the determinants of scientific production at cross-country level. The aim of this paper is not to provide accurate and robust estimates of investment elasticities (a doubtful task given the poor quality of the data sources and the modelling problems), but to develop and critically assess the validity of an empirical approach for characterising the production of science and its impact from a comparative perspective. We employ and discuss the limitations of a production function approach to relate investment inputs to scientific outputs using a sample of 14 countries for which we have information about Higher Education Research and Development (HERD). The outputs are taken from the Thomson ISI® National Science Indicators (2002) database on published papers and citations. The inputs and outputs for this sample of countries have been recorded for a period of 21 years (1981-2002). A thorough discussion of data shortcomings is presented in this paper. On the basis of this panel dataset we investigate the profile of the time lag between the investment in HERD and the research output and the returns to national investment in science. We devote particular attention to analysis of the presence of cross-country spillovers. We show their relevance and underline the international effect of the US system

Keywords: Productivity of science, lag structure, returns to HERD investment, international spillovers.

JEL Subject Classification: L3, O3

1 Introduction

There is increasing recognition in the OECD countries of the importance of public scientific research in providing the foundations for both innovation and competitiveness. At the same time, there is little systematic evidence on how such investments can lead to increasing levels of scientific output, improved patenting and innovative output, better economic performance and, ultimately, to increased wealth for a country. Much of the available literature concentrates either on investigation of the effects of public basic research on the innovative activities of firms (see among others Jaffe, 1989; Mansfield, 1991; Narin et al., 1997; Klevorick et al., 1995; Arundel and Geuna, 2004) or the contribution of scientific research to productivity growth (Adams, 1990).[1] Only a very few studies have focused on the relationship between investment in science, and measures of research outputs, discussed below in detail (of particular interest are the works of Johnes and Johnes (1995), Adams and Griliches (1998) and Bonaccorsi and Daraio (2003).

The need for justifying expenditure on science relative to other claims on public expenditure have influenced the method and focus of analyses of the economic contribution of public scientific research. Though acknowledging the limitations of the 'linear model,' in which new technology is the inevitable consequence and ultimate benefit of funding basic research investment, the push for quantification and measurement has constrained most of the studies to implicitly represent the system of 'payoffs' of science investment as being a set of connected black boxes starting with the investment in public scientific research and ending with the creation of wealth or production of socio-economic benefits (see Figure 1). Although Figure 1 does depict the first-order feedbacks between actors, there is a much richer and more complex set of interactions among the four major actors. Towards the right end of the figure there are heroic assumptions allowing the connection between inputs such as firm innovation and patents and sectoral outputs and productivity performance, the principal determinants of wealth creation (producer and consumer surplus). Feedbacks, external factors, firm heterogeneity, industrial sectors knowledge bases variability, etc. are not included in the estimations of the socio-economic impact of the investment in public scientific research. Paradoxically, most of the literature has focused on the second (firms) and third (sectors) boxes, while little analysis of the first (universities) has been conducted despite the considerable amount of data available for such analysis.

{FIGURE 1 ABOUT HERE}

The aim of this study is to contribute to the development of a better understanding of the relationships between government R&D funding and scientific production using a production function approach. In so doing, we will examine the limits in this approach that arise from the fact that the inputs to scientific production as well as its outputs are

[1] In recent years very many papers have analysed university-industry relationships and university technology transfer and intellectual property rights developments (for an overview of the main trends see OECD, 2002a, 2003a; Geuna and Nesta, 2005; Mowery et al., 2004). Most of these works have tended to be focused on the characteristics of the actors involved and not the relationships between inputs and outputs in the long and recursive process between science investment and wealth creation. The contribution of public research spending is sometimes examined within broader analyses of productivity growth or economic growth of countries (Guellec and van Pottelsberghe, 2001; OECD, 2003b). For a review of the literature on the economic return to public scientific research see Scott et al. (2001) and Steinmueller (2001).

difficult to quantify and price. The analysis is undertaken from an international viewpoint, considering that there is a 'world' science production function – an approach that assumes a relatively efficient and free flow of scientific knowledge at the global level. We identify (not impose) the structure and length of the lag between investment in R&D and scientific outputs. Finally, we analyse international spillovers, and the impact of the US system on other countries' science systems.

We employ a sample of 14 OECD countries for which we have reliable information about Higher Education Research and Development (HERD). As proxies for outputs we use publications and citations from the Thomson ISI(R) National Science Indicators (2002) database. The inputs and outputs for this sample of countries have been recorded over a period of 21 years (1981-2002). On the basis of this panel dataset we develop an empirical framework to address the following research questions:

i. What is the profile of the time lag between investment in HERD and research outputs? Does it vary depending on the output being considered?
ii. What are the returns to national investment in science? Are there cross-country spillovers?

The paper is organised as follows. Section 2 presents the model and data, and discusses their shortcomings. Section 3 presents a preliminary econometric estimation of the lag length and structure between science funding and scientific research outputs at the international level. International spillovers and the final version of our model are described in Section 4. Finally, Section 5 concludes, and also suggests some avenues for further research.

2 Modelling and measuring scientific production

2.1 A knowledge production function model

Following Griliches (1979), a simple knowledge production function can be specified as follows. Let $Y = F(X, K, u)$ be the production function connecting some measure of output, Y, at the micro (researcher) or macro (country) level, to the 'inputs' X, K, and u; where X stands for an index of conventional inputs such as labour and other control variables, K is a measure of the current state of scientific knowledge, determined in part by current and past research and development expenditures and u represents all other unmeasured determinants of output and productivity.

A common approach is to specify the production function as a Cobb-Douglas and assume that the unmeasured factors u can be considered as random after the introduction into the equation of a time trend to represent the systematic component of the unmeasured factors. Then we can rewrite F(.) as:

$$Y_{it} = A_i X_{it}^{\alpha} K_{it}^{\beta} e^{\lambda t + u_{it}} \quad (1)$$

where A is a constant at the country level, i is a country index, t is a time index, e is the natural logarithm base and α, β and λ are the parameters that we are interested in estimating. The greatest controversy arises around the specification of the knowledge capital K_{it}. According to Griliches (1979), three major issues should be considered in its measurement: (a) the fact that the research process takes time and that current research and development may not have an effect on measured productivity until several years have elapsed; (b) past research and development investments depreciate and become obsolete and therefore growth in the 'net' stock of knowledge capital is not equal to the gross level of current or recent resources invested to expand it; and (c) that the level of knowledge for a given research unit (or country) is not only derived from 'own' research and development investments, but is also affected by the knowledge of other units (or countries) through knowledge spillovers.

Regarding the time lag, it has been traditional to assume a linear lag polynomial specification which is consistent with an immediate 'impulse' to output followed by a declining effect. An alternative to the linear model, which implies perfect *substitution* between research and development expenditures carried out over different periods of time, is to assume that the old capital and the new investment are *complementary* inputs in the production of new knowledge capital (Klette and Johansen, 2000). The basic idea is that the greater the initial knowledge, the greater will be the amount of knowledge obtained from a given amount of R&D. This assumption is particularly apt in the case of science where we 'stand on the shoulders of giants' to build new knowledge. The more knowledge has been produced, the more it can be recombined to produce new knowledge. Formally we will assume that:

$$K_{it} = \prod_{j=0}^{\infty} R_{it-j}^{w_j} \quad (2)$$

Equation (2) has also the advantage that it makes the model estimation linear (in logs) allowing us to search for rather than to impose the pattern of weights (w) in the lag structure.

The issue of the rate of depreciation of knowledge capital is more complex. While it is clear that private knowledge capital sometime depreciates rapidly (when products and process from competitors reach the market), much less is known about the impact of depreciation on the 'public' stock of knowledge. Such knowledge may remain useful for quite long periods of time, and hence be subject to much lower rates of depreciation. In the estimations we will assume that the lag length is finite, but quite extensive.

The world of 'open' science rewards scientists for speedy disclosure of their discoveries, an incentive that is antithetical to keeping a discovery secret in order to appropriate the possible economic returns from its exploitation.[2] International collaborations are more common in scientific developments than for industrial developments. In this collaborative environment knowledge spillovers should be pervasive. The science production function (1) should be corrected to include a cross-country knowledge spillover term, whose estimation we pay particular attention to.

Finally, there is the issue of aggregation problems (Fisher, 1969; Griliches, 1979; Felipe and Fisher, 2003). Traditional estimations of industrial production function at

[2] This view of the 'open' science organisation system for the production of new knowledge (Dasgupta and David, 1994) is currently being challenged by a more proprietarily oriented model based on university property rights (see for example Mowery et al., 2004 for a discussion of the US situation and Geuna and Nesta, 2005 for a discussion of the EU).

country level are affected by the way in which inputs and outputs are aggregated (the major problems usually are centred around inputs). We assume a situation of equilibrium at the micro level that allows one to sum-up to the sectoral and then country level using prices. In the case of a science production function the research outputs (publications and citations) are an aggregation of a very diverse set of items: different publications, in different journals, in different scientific fields, with the different propensity to publish and different propensity to produce journal publications as their codified output. Clearly we do not have 'prices' that could permit us to sum across the various categories in a homogeneous way. Does this mean that we cannot estimate a production function at the macro level for science?

We can express our output indicator as:

$$Y_{it} = \prod_{l=1}^{L} y_{ilt}^{\Omega_l} = A_i X_{it}^{\alpha} K_{it}^{\beta} e^{\lambda t + u_{it}} \quad (3)$$

where the 'aggregate' output Y_{it} is given by a number index form with weights Ω_l of some *unobserved* indicators of the quality of each type of publication output. We assume that there are L different types of research outputs. Because we do not observe Ω_l, we have assumed a common weight for each type of publication (or scientific field). The deviations of the true weight regarding the average weight will appear as $\sum_{l=1}^{L} \varepsilon_{ilt} y_{ilt}$ in the residual of the knowledge production function as in (4)

$$\sum_{l=1}^{L} \overline{\Omega} y_{ijt} = a_i + \alpha x_{it} + \beta k_{it} + \lambda t - \sum_{l=1}^{L} \varepsilon_{ilt} y_{ilt} + u_{it} \quad (4)$$

We can assume that the within country set of weights remains stable over time, that is, we assume that there is no major change in the type or frequency of publications in a country (or that if there is it is similar across countries).[3] However, the size of the countries' research outputs Y_{it} influences the estimation. If the relative differences in the scientific research size among countries remain constant over time (an acceptable assumption for this sample of countries given their highly developed science system), it would be possible to absorb this omitted factor within a country specific fixed effect. Unfortunately, the lack of sufficiently detailed comparable data precluded us from directly testing this assumption. However, in order to further study if our results are affected by scale effects, we tested for the stability of the knowledge production function across different sub-samples of countries, and found the results to be stable.[4]

2.2 Data sources

We focused our analysis on 14 countries: Australia, Belgium, Canada, Finland, Denmark, France, Germany, Netherlands, Spain, Italy, Switzerland, Sweden, the UK

[3] This assumption constrains the analysis to countries with similar levels of science development, so that we can assume that a change, due for example to the arrival of genomics, affects the portfolio of research output of the various countries in similar ways. This model would not be robust to the inclusion of developing countries in the sample.
[4] The results of this test are available from the authors on request.

and the USA. We excluded other OECD countries where information about HERD was incomplete and/or inconsistent, and those countries that had a specific scientific research output (due to their size, history or other factors) and therefore were causing problems with the aggregation of the outputs.

In order to examine the relationship between science funding and scientific research outputs we used the publicly available information on HERD expenditure and its components at country level. The OECD defines the HE sector as universities, colleges of technology and other institutions of post-secondary education, whatever their source of finance or legal status. This includes research institutes, experimental stations and clinics operating under the direct control of, administered by or associated with the higher education institution (HEI) (OECD, 2002b). Because this sector does not usually directly match with an area in the System of National Accounts, it is difficult to provide clear guidelines that ensure internationally comparable data reporting. Universities and colleges of technology make up the core of the sector in all countries. Variations occur with respect to other post-secondary education institutions and even more so with respect to institutes linked to universities, such as university hospitals and clinics and to public research centres - e.g. CNRS laboratories are included in HERD in France while similar institutions in Italy, i.e. CNR laboratories, are included in the 'government' research and development expenditure category). Also, country reporting differs in how HE expenditure is classified - e.g. PhD students in Sweden and The Netherlands receive a state salary (they were or are public employees depending on the time period considered), while those in the UK and Italy receive only a grant (Geuna, 2001; Jacobsson and Rickne, 2003). These differences limit the validity of cross-country comparisons. A case-by-case analysis was carried out in order first to identify 'major' structural breaks in the series and second to select a set of comparable countries with comparable statistics.[5] However, because 'permanent' country level differences in the way that information is collected persist, we had to control for these systematic differences in the estimations.[6]

The HERD figures that we used are expressed in millions of constant US$ as reported by the OECD. R&D expenditure series were deflated using the implicit gross domestic product (GDP) deflator taken from the OECD National Accounts database. These national currency data at 1995 prices were converted to US$ using 1995 purchasing power parities (PPP).[7]

---

[5] In particular all the models presented in this paper were first estimated on a country-by-country basis and then different dummy variables were included if there was some report about changes in the classification criteria. Only those countries for which results were robust to these breaks were considered.

[6] More specifically, a country level fixed effect was always introduced into the models. This fixed effect not only captures permanent differences relating to the functioning of the various national scientific systems, but also differences in how the information is collected. More particularly, for almost all the results shown below we worked with 'within' country information and 'averaged' to achieve a global estimation. We were able to deploy this approach due to the panel data nature of the dataset we built.

[7] The fact that we are using the GDP deflator instead of the HERD deflator provides another interpretation of the time index. For example, if the HERD deflator is increasing more rapidly than the GDP deflator (the most likely relationship), the time index will reflect this effect – offsetting positive contributions to the productivity of research inputs arising from exogenous and unmeasured effects or enhancing negative effects such as increases in the bureaucratic requirements required to justify and account for research funding (or both).

The scientific process produces several research outputs that can be classified into three broadly defined categories: (1) new knowledge; (2) highly qualified human resources; and (3) new technologies. This paper focuses on the determinants of the first type of research output.[8] There are no direct measures of new knowledge, but various proxies have been used in previous studies. The most commonly used are (a) publications and (b) citations. The source of these two variables is the Thomson ISI® 'National Science Indicators (2002) database of published papers and citations. These two measures have several shortcomings (Geuna, 1999); here we detail only those that directly affect the econometric estimations. First, they are incomplete and biased proxies for the production of new knowledge. Second, the Thomson ISI data are strongly affected by the disciplinary propensity to publish in journals, so they are a poor measure of the output of disciplines such as history or law. Third, the ISI includes an almost constant number of journals/pages in its archive (journals enter and leave, but the number remains more or less constant at around 5,000; journals can increase the number of issues per year, but this happens only for a minority of journals). This clearly creates limits the possibility of output expansion and therefore biases our estimations in favour of decreasing returns.

In order to take account of the 'truncation problem' in the citations to most recent years, the citations variable has been adjusted.[9] One way of controlling for this is to use what Hall et al. (2001) called the fixed effect approach. This method involves scaling citations counts by dividing them by the average citation count for a group of papers to which the paper in question belongs. Using the same example as in Hall et al., this approach treats a paper that received say 11 citations and belongs to a group in which the average paper received 10 citations, as equivalent to a paper that received 22 citations, but happens to belong to a group in which the average was 20. The groups were defined in terms of scientific field and year, and the scaling index was computed using the ISI dataset at world level.

Both publications (and citations) and HERD expenditure levels have a low variance, i.e. they tend to grow steadily over time. This greatly complicates the statistical inference because these series can be correlated due to the presence of a third common variable (the time trend) leading to spurious or overestimated results. In order to determine what kind of stochastic process governs the series, we carried out several tests. Panel Data Unit root tests were used to determine whether it is better to work with first differenced observations (because they have a unit root) or to control by a time trend in the models (because they are a trend stationary series). While the results of these tests are omitted here, they provide a basis for treating the series as stationary, but with a deterministic trend. In what follows we work with these series in levels and add a deterministic trend to the models.[10]

[8] Preliminary estimates of the determinants of the second type of research output (highly qualified human resources) can be found in Crespi and Geuna (2004).

[9] The citation count is affected by the time span allowed for the papers to be cited: for example, papers published in 2000 can receive citations in our data just from papers published in the period 2000-2001; in fact they will be cited by papers in subsequent years as well, but we do not observe these.

[10] For further details about the results of these tests see Appendix B in Crespi and Geuna (2004).

Acknowledging these input and output data limitations, as well as the modelling problems and strong assumptions discussed in the previous section, we advise caution in interpreting the results of the econometric models. In particular, it is inappropriate to conclude that estimates of the investment elasticities are accurate or robust (a doubtful task given the poor quality of the data sources and the modelling problems). Our aim is to identify some basic facts about the operation of the production of science and its impact for which issues of the delay of impacts and the role of spillovers are of central concern. In other words, the aggregate science production function is used in this context not as a tool for accounting for output at a national level, but as an instrument for highlighting the structural characteristics of the process of scientific production.

In the next section we present the econometric model used to identify the structure and length of the lag between investment in HERD and HE outputs. Once the lag structure was identified we focused on the search for international spillovers. Section 4 presents the final estimation of our model (see Table 6), which includes both national and international HERD expenditures.

3 The Polynomial Distributed Lag Model

One (but not the only) way to search for both lag length and structure is to apply the technique known as Polynomial Distributed Lag (PDL) or the Almon Model (see Greene, 1993). This methodology can be applied to our case. Let us define the following 'finite' distributed lag model:

$$y_{it} = \alpha_i + \sum_{j=0}^{q} \beta_j r_{it-j} + \gamma_i X_{it} + u_{it}, \quad i = 1,.....N \qquad (1)$$

where y_{it} is the log of given research output (publications and citations) and r_{it} is the log of HERD, for country i at time t. Although a model like (1) in theory can be estimated quite straightforwardly, there is a potential problem of very long lags in which case the multicolinearity is likely to become quite severe. In such examples it is common to impose some structure on the lag distribution, reducing the number of parameters in the model. It is in this context that the PDL model can be useful. The approach is based on the assumption that the true distribution of the lag coefficients can be very well approximated for by a polynomial of a fairly low order.

$$\beta_j = \delta_0 + \delta_1 j + \delta_2 j^2 + ...+ \delta_p j^p, \quad j = 0,.....q > p \qquad (2)$$

The order of the polynomial, p, is usually taken to be quite low, rarely exceeding 3 or 4. By inserting (2) into (1), a transformed model can be estimated, in which the estimated coefficients are deltas that can be put back into (2) in order to recover the original weights. In addition to the $p+1$ parameters of the polynomial, there are two unknowns to be determined: the length of the lag structure, q, and the degree of the polynomial, p. Here we follow the standard procedure for determining first the lag lengths and then the degree of the polynomial function.

Table 1 presents the results for publications and Table 2 the results for citations. Both tables show the lag structure for each alternative model and in the last three rows present the values for the information criteria and the long run elasticity for domestic HERD. In the results for publications (Table 2.1) out of the three information criteria

AIC, SIC and ICOMP, two have a maximum value at a 5-year lag, while the remaining criterion reaches a minimum at a 6-year lag. Because of the potentially more serious consequences of omitting some relevant lag, we decided to keep 6-year lags as the optimum lag length for publications.

To search for lag length we started by taking a lag of 10 years and reducing by one period down to 0. In each reduction we evaluated the information that is lost with the omission of one additional lag, against the information that is gained as a result of having more degrees of freedom in the estimation.[11]

Table 1: Unrestricted Polynomial Distributed Lag (PDL) Model Approach (Fixed Effects) Publications

	10	9	8	7	6	5	4	3	2	1	0
	Lpub	Lpub	Lpub	Lpub	Lpub	Lpub	Lpub	Lpub	Lpub	Lpub	Lpub
HERD	-0.104	-0.091	-0.085	-0.082	-0.076	-0.080	-0.104	-0.140	-0.157	-0.288	0.012
	0.121	0.119	0.116	0.115	0.114	0.114	0.116	0.124	0.132	0.137**	0.092
t-1	0.073	0.065	0.068	0.069	0.062	0.060	0.054	0.068	-0.009	0.366	
	0.145	0.146	0.145	0.145	0.144	0.146	0.151	0.157	0.158	0.140***	
t-2	0.030	0.040	0.041	0.039	0.043	0.040	0.052	0.000	0.326		
	0.138	0.139	0.139	0.138	0.138	0.139	0.145	0.140	0.115***		
t-3	0.103	0.103	0.098	0.098	0.094	0.097	0.043	0.306			
	0.131	0.131	0.131	0.131	0.130	0.131	0.132	0.102***			
t-4	0.046	0.043	0.047	0.045	0.051	0.021	0.265				
	0.117	0.117	0.117	0.116	0.116	0.112	0.084***				
t-5	0.113	0.114	0.107	0.109	0.092	0.236					
	0.104	0.104	0.103	0.102	0.102	0.073***					
t-6	0.073	0.067	0.073	0.067	0.145						
	0.082	0.083	0.082	0.083	0.058**						
t-7	0.065	0.070	0.051	0.079							
	0.070	0.071	0.072	0.057							
t-8	-0.029	-0.047	0.028								
	0.078	0.079	0.053								
t-9	0.006	0.076									
	0.104	0.067									
t-10	0.073										
	0.085										
Non-HERD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	-0.001	-0.001	-0.001
	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.005	0.005
Year	0.015	0.016	0.017	0.017	0.018	0.020	0.023	0.027	0.030	0.034	0.036
	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.003***
Constant	-23.419	-24.887	-26.404	-27.013	-28.829	-32.605	-38.990	-45.803	-52.081	-57.875	-61.893
	7.865***	7.680***	7.603***	7.442***	7.349***	7.246***	7.105***	7.203***	6.829***	6.468***	6.153***
Observations	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000
AIC	-428.68	-429.89	-430.97	-432.85	-433.81	-431.96	-424.1	-414.71	-403.33	-395.26	-386.47
SIC	-347.46	-351.79	-356	-360.99	-365.08	-366.36	-361.62	-355.35	-347.1	-342.15	-336.48
ICOMP	NA	-262.66	NA	-274.57	-280.22	-283.61	-281.24	-276.13	-271.02	-268.05	-262.88
HERD_LR	0.45	0.44	0.43	0.42	0.41	0.37	0.31	0.23	0.16	0.08	0.01
	0.098***	0.096***	0.096***	0.096***	0.096***	0.098***	0.099***	0.104**	0.100	0.096	0.092

Robust standard errors reported below each coefficient.
(*) significant at 10%; (**) significant at 5%; (***) significant at 1%

[11] We use three different statistics here: the Akaike Information Criteria (AIC), the Schwartz Bayesian Information Criteria (SBC) and the Bozdogan index of Information Complexity (ICOMP). These criteria are used to select the 'best' model by balancing an adequate goodness of fit against a small number of parameters. See Kolenikov (2000) for the source codes for STATA.

In our estimation we used two control variables: a time trend to capture the evolution of the general scientific opportunity and a variable capturing the proportion of non-HERD R&D in the total country research budget. The rationale for including the latter variable derives from the fact that our 'observed' research output is total (not just HE) country-level publications and citations. Even though most (over 80%) of the publications generated by a country are typically derived from the research being carried out in HE institutions, there is still a small proportion produced by firms and other non-university research centres. We would expect the productivity of non-HERD institutions in terms of publications and citations to be lower than that of universities. Publications and citations are a by-product of the innovation activities or research supporting government actions of these non-HERD institutions. An increased proportion of non-HERD R&D in total Gross Expenditures in Research and Development (GERD) would lead to a reduction in a country's total publications and citations. In order to control for this, we built a new variable defined as the ratio between non-HERD R&D and GERD (*Non-HERD*).

Table 2: Unrestricted Polynomial Distributed Lag (PDL) Model Approach (Fixed Effects) Citations

Lag Num	10	9	8	7	6	5	4	3	2	1	0
	Lcit	lcit	Lcit	Lcit	Lcit	Lcit	lcit	Lcit	lcit	lcit	Lcit
HERD	-0.269	-0.257	-0.251	-0.243	-0.229	-0.237	-0.271	-0.328	-0.354	-0.546	-0.129
	0.166	0.165	0.162	0.161	0.156	0.156	0.159*	0.168*	0.178**	0.179***	0.115
t-1	0.092	0.084	0.087	0.089	0.074	0.070	0.060	0.082	-0.041	0.508	
	0.206	0.205	0.205	0.204	0.200	0.202	0.213	0.226	0.232	0.198**	
t-2	0.013	0.022	0.023	0.017	0.027	0.021	0.038	-0.041	0.477		
	0.168	0.167	0.166	0.166	0.166	0.167	0.179	0.176	0.161***		
t-3	0.170	0.171	0.165	0.166	0.156	0.162	0.082	0.485			
	0.158	0.158	0.158	0.158	0.157	0.159	0.161	0.124***			
t-4	0.088	0.085	0.089	0.086	0.097	0.044	0.407				
	0.133	0.133	0.132	0.131	0.132	0.127	0.105***				
t-5	0.129	0.130	0.123	0.128	0.089	0.352					
	0.118	0.118	0.117	0.116	0.121	0.085***					
t-6	0.102	0.096	0.102	0.084	0.264						
	0.101	0.103	0.102	0.104	0.075***						
t-7	0.114	0.119	0.100	0.182							
	0.090	0.092	0.093	0.075**							
t-8	0.025	0.008	0.084								
	0.102	0.104	0.069								
t-9	0.011	0.077									
	0.111	0.079									
t-10	0.070										
	0.086										
Non-HERD	-0.008	-0.008	-0.008	-0.008	-0.008	-0.008	-0.009	-0.010	-0.009	-0.010	-0.010
	0.005*	0.005*	0.005*	0.005*	0.005*	0.005	0.005*	0.005*	0.006*	0.006	0.006
Year	0.013	0.013	0.014	0.015	0.017	0.021	0.026	0.032	0.037	0.042	0.045
	0.005**	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***	0.004***	0.004***
Constant	-18.836	-20.240	-21.787	-23.624	-27.831	-34.710	-44.219	-54.682	-64.647	-73.110	-78.685
	9.650*	9.161**	9.107**	8.993***	8.882***	9.005***	8.709***	8.822***	8.451***	7.802***	7.625***
Observations	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000	168.000
AIC	-385.630	-387.070	-388.330	-389.450	-387.190	-379.670	-365.950	-349.340	-329.110	-317.450	-306.450
SIC	-304.410	-308.970	-313.360	-317.600	-318.470	-314.070	-303.460	-289.980	-272.880	-264.340	-256.470
ICOMP	-211.910	-219.300	-224.440	-230.040	-232.810	-229.720	-222.080	-210.100	-196.660	-190.980	-182.590
HERD_LR	0.543	0.533	0.522	0.508	0.477	0.411	0.316	0.198	0.082	-0.037	-0.128
	0.125***	0.121***	0.121***	0.121***	0.122***	0.130***	0.131**	0.139	0.133	0.122	0.115

Robust standard errors reported below each coefficient.
(*) significant at 10%; (**) significant at 5%; (***) significant at 1%

Table 2 summarises the results for citations. According to one of the criterion noted in footnote 11, the optimum lag length is 7 years, while based on the other two it is 6. Again, we chose the conservative approach and used the longer lag.

It is interesting to compare the results for publications and citations in terms of the long-run elasticities implied by the sum of all the individual coefficients. In the case of publications, the long-run elasticity is 0.41 and is statistically significant, and for citations it is 0.51 and also significant. In both cases the long run elasticities reported above become quite stable to small variations in the lag length. In addition, for both publications and citations we find that the Non-HERD share variable is negative although significant only for citations, while the time trend is positive and significant for both outputs.

Table 3: Unrestricted PDL and Restricted Almon Models (Fixed Effects)

	Unrestricted	Restricted	Unrestricted	Restricted
	Publications	Publications	Citations	Citations
HERD	-0.076	-0.017	-0.243	-0.032
	0.113	0.018	0.161	0.017*
t-1	0.062	-0.001	0.089	-0.021
	0.144	0.033	0.202	0.024
t-2	0.043	0.033	0.017	0.017
	0.139	0.020*	0.170	0.022
t-3	0.094	0.078	0.166	0.073
	0.131	0.013***	0.158	0.016***
t-4	0.051	0.109	0.086	0.124
	0.116	0.013***	0.130	0.012***
t-5	0.092	0.117	0.128	0.157
	0.102	0.017***	0.116	0.015***
t-6	0.145	0.085	0.084	0.158
	0.058**	0.015***	0.104	0.017***
t-7			0.182	0.111
			0.074**	0.014***
Non-HERD	0.000	0.000	-0.008	-0.006
	0.001	0.001	0.004*	0.003*
Year	0.018	0.018	0.015	0.012
	0.004***	0.003***	0.004***	0.003***
Constant	-28.829	-29.505	-23.624	-18.009
	7.354***	5.808***	8.982***	6.795***
Observations	168	168	168	168
HERD_LR	0.410	0.405	0.508	0.587
	0.096***	0.071***	0.121***	0.081***
Constraints (Chi2)		1.54		1.06
Polynomial Degree		Critical		Critical
5 to 4	1.26	6.63	3.710	6.63
4 to 3	1.45	7.83	4.580	7.83
3 to 2	12.53**	8.97	27.87***	8.97
2 to 1	59.27***	10.06	98.05***	10.06

Robust standard errors reported below each coefficient.
(*) significant at 10%; (**) significant at 5%; (***) significant at 1%

Assuming that we have identified the appropriate lag length we then need to identify the right polynomial function. We proceeded by using a 5th degree function and testing sequential unit reductions in the degree. The results are shown in the last 4 rows of Table 3, where we can accept the reduction from 5th to 4th and from 4th to 3rd but not lower. It is important to note that in order to keep the appropriate significance level in each step we used a very low individual significance level. The choice of a 3rd degree polynomial function is therefore appropriate for both publications and citations.

The PDL model also implies a set of constraints on the unrestricted model (without a specified functional form for the lags) estimated above. For example, in the case of publications the optimum lag length is 6 and we use a 3rd degree polynomial function, we are implicitly imposing three constraints. In addition there are endpoint constraints that allow the lag distribution to be 'tied down' at its extremes. These endpoint constraints capture the idea that there is no effect from R&D on the research outputs *before* the current period[12] and also that there is no effect from the research inputs after the maximum lag. That is, we need to impose:

$$\beta_{-1} = 0 \qquad and \quad \beta_{q+1} = 0 \qquad (3)$$

In total we have five constraints. One way of validating the PDL model is to test whether these constraints are valid. As shown by the non significant Chi test in Table 3, we could not reject any of them. In terms of long-run elasticities we found that their values are very similar to those for the unrestricted model for publications and slightly higher for citations. The time trend is positive and significant, while the non-HERD institutions have a negative effect only on citations.

It is important to compare the weights for the unrestricted model with those obtained using the restricted model. The pattern for both publications and citations is similar. The impact in the first two years is always very low and not significantly different from zero. It is only at the end of the 2nd year for publications (3rd year for citations) that there is a positive impact which peaks at year 5 for publications (year 6 for citations).

Figures 2 and 3 are graphical representations of the lag structures implied by the restricted (dotted line) and the unrestricted models. These figures clearly show the positive impact from the investment in science at the end of year 2 for publications and year 3 for citations, with increasing returns till the end of year 5 (or year 6 in the case of citations) and subsequent decline till the end of year 6 (or year 7 for citations). Apparently, after six (seven) years no significant returns can be expected.

[12] This means that the research output does not react 'in advance' of an increase in the research inputs.

Figure 2. Unrestricted vs Restricted Pattern of Weights (Publications)

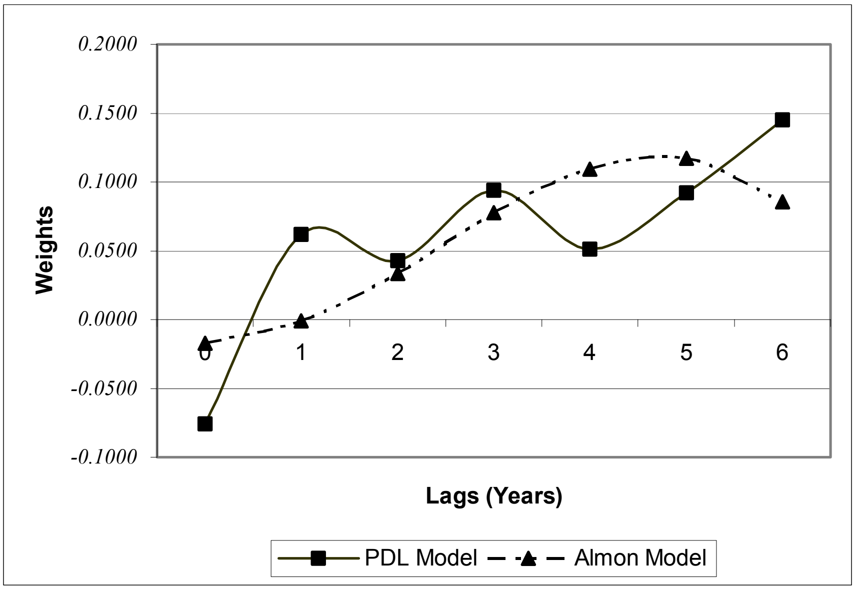


Figure 3. Unrestricted vs Restricted Pattern of Weights (Citations)

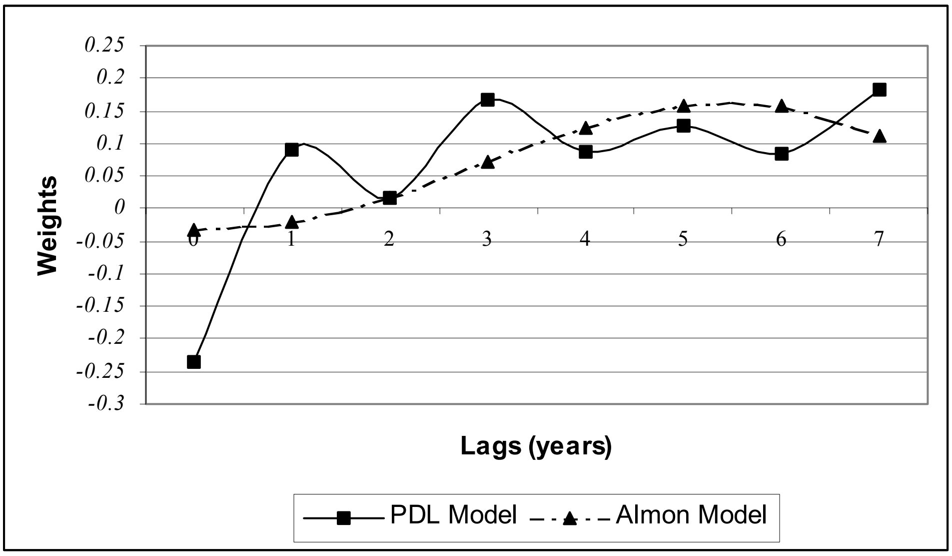
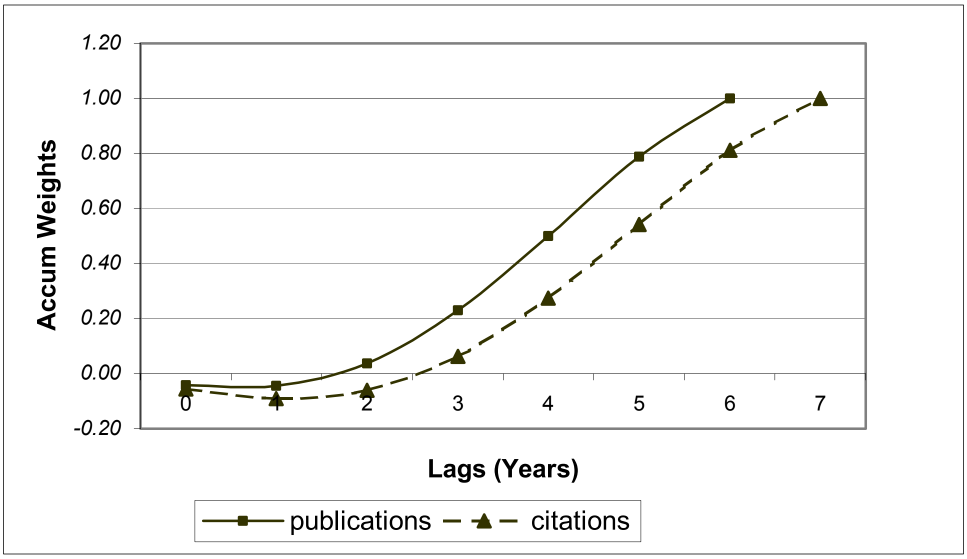


Figure 4 shows the evolution of the cumulative impact of HERD expenditure on scientific research outputs. There is no positive 'cumulative' impact until year 2 in the case of publications (year 3 in the case of citations) and it is not until year 4 (or year 5 in the case of citations) that 50% of the expected impact occurs.

Figure 4. The Cumulative Patterns of Weights (Publications and Citations)



The PDL model makes use of the lag structure of (the log of) past R&D. This implies a form of Cobb-Douglas knowledge-creation function where there is unit substitution elasticity between current and past R&D expenditure. It is important to say that this kind of function is slightly different from the traditional one, which is assumed by Adams and Griliches (1998) when they use the log of the weighted sum of past R&D. This implies a linear knowledge-creation function where there is a sort of perfect substitution between current and past R&D expenditure. Working with this sort of function in the context of panel data is very complex: to cancel out the fixed effects by using either within or first difference transformations when the underlying function is not linear is far from straightforward. Also, simply adding country dummies to the model complicates the non-linear estimation process. Having said this, we managed to calculate the non-linear estimate for the lag structure for publications. Both the optimum lag structure and the profile of the weights were similar to our original linear estimations. We therefore decided to proceed with the simpler linear model.

4 The Search for Spillovers

We investigated whether spillovers occur between countries. In this context, 'spillovers' means that part of the increase in the research output of a given country is due to HERD investments made in other countries.

To the extent that an ever growing external pool of knowledge available to each country generates these spillovers, the inclusion of a time-trend variable in the model will partially capture them. Thus, the inclusion of a specific spillovers variable should not to any significant extent affect the estimated domestic HERD elasticity. However, identification of a specific spillover effect allows us to calculate the total (domestic plus international) return from changes to the HERD investment for the system of countries considered.

In order to estimate the existence of spillovers, we need to assess the level of knowledge exchange or knowledge cooperation among countries. The higher the level of exchange/cooperation between countries *i* and *l*, the higher is the probability that some of the science investment in country *i* will affect the research output in country *l* and vice versa. To build up a matrix of knowledge proximity among countries, we used the information on international scientific co-authorship. The National Science Foundation's *Science and Engineering Indicators* reports (NSF, various years) give the share of cross-country co-authorships through the 1980s and 1990s. We averaged the values and built a weight for each country as follows:

> w_{il} = number of international co-authorships between countries *i* and *l*, divided by the total number of international co-authorships between country *l* and other countries in the sample.

This weight provides a proxy for the relative knowledge exchange/cooperation between two given countries in our dataset. Table 4 presents the resulting weights for the 14 countries in our database. The US is the most important collaboration partner for all the countries considered (the US always has the highest weight). This indicates the important role of the US in the process of knowledge creation. Table 4 also clearly indicates how geographical proximity and cultural and linguistic links, such as apply to Belgium and France, and the UK and Australia, affect co-authorship patterns.

Table 4: Weighting Matrix

	AU	B	CA	Dk	Fin	F	G	I	NL	E	S	CH	UK	US
AU	0.000	0.012	0.087	0.017	0.007	0.051	0.088	0.025	0.028	0.008	0.031	0.024	0.214	0.409
B	0.011	0.000	0.035	0.015	0.014	0.192	0.118	0.065	0.120	0.035	0.036	0.050	0.107	0.201
CA	0.036	0.014	0.000	0.013	0.009	0.090	0.061	0.030	0.026	0.012	0.022	0.025	0.108	0.552
Dk	0.021	0.019	0.040	0.000	0.033	0.070	0.129	0.056	0.048	0.029	0.136	0.040	0.137	0.242
Fin	0.015	0.027	0.043	0.049	0.000	0.062	0.123	0.047	0.053	0.019	0.148	0.046	0.098	0.271
F	0.016	0.061	0.069	0.017	0.011	0.000	0.135	0.094	0.045	0.058	0.030	0.073	0.117	0.275
G	0.024	0.032	0.038	0.027	0.017	0.115	0.000	0.069	0.060	0.031	0.041	0.090	0.121	0.334
I	0.012	0.031	0.033	0.020	0.011	0.140	0.122	0.000	0.044	0.044	0.033	0.082	0.130	0.298
NL	0.019	0.080	0.041	0.025	0.018	0.094	0.148	0.060	0.000	0.029	0.036	0.049	0.147	0.255
E	0.007	0.036	0.028	0.021	0.009	0.186	0.114	0.093	0.043	0.000	0.024	0.040	0.156	0.241
S	0.024	0.029	0.040	0.083	0.061	0.075	0.118	0.053	0.042	0.020	0.000	0.042	0.114	0.300
CH	0.014	0.030	0.035	0.019	0.015	0.138	0.199	0.103	0.044	0.024	0.032	0.000	0.099	0.248
UK	0.056	0.029	0.068	0.028	0.014	0.098	0.118	0.072	0.059	0.041	0.039	0.044	0.000	0.335
US	0.053	0.026	0.171	0.024	0.018	0.113	0.162	0.082	0.050	0.032	0.050	0.055	0.165	0.000

Source: Weighting Matrix. Author's own elaboration based on NSF data.
Australia (AU), Belgium (B), Canada (CA), Denmark (Dk), Finland (Fin), France (F), Germany (G), Italy (I), Netherlands (NL), Spain (E), Sweden (S), Switzerland (CH), United Kingdom (UK) and United States (US).

After building these weights we defined the 'international' research and development relevant to each country as a weighted sum of the science budgets of all the other countries as follows:

$$S_{lt} = \sum_{i \neq l} w_{il} R_{it} \quad (5)$$

where R_{it} is the HERD budget for country *i*. After constructing (5) we assume that the lag structure is the same as was found in Section 3.[13]

The model estimated in relation to spillovers focuses on long-run spillover effects. In order to compare these effects with the long-run impact of domestic R&D, we redefined the stock of knowledge as the weighted sum of (the log of) R&D, where the weights are defined (as in Green, 1993) as follows:

$$\omega_j = \frac{\beta_j}{\sum_{j=1}^{k} \beta_j} \qquad (6)$$

and we used the weight ω_j in order to aggregate the lag for (the log of) R&D expenditure for each country in the dataset. In this way we dispense with the need to estimate short-run elasticities (which we assume to be known) and instead focus on long-run elasticities.

Table 5: Results using 6 (7) Lags of RD for Publications (Citations) plus Spillovers [dependent variable log Publications (Citations)] Fixed Effects by Country Included

	Publications (1)	Publications (2)	Citations (3)	Citations (4)
Non-HERD$_{it}$	-0.005	-0.004	-0.013	-0.012
	0.001***	0.002**	0.002***	0.002***
Year	0.018	-0.000	0.014	-0.014
	0.002***	0.0020	0.003***	0.006**
HERD$_{it}$	0.475	0.447	0.536	0.499
	0.047***	0.045***	0.049***	0.047***
S$_{it}$		0.505		0.599
		0.116***		0.123***
Constant	-35.024	-2.859	-21.162	27.627
	4.444***	7.9410	6.187***	11.60**
Observations	224	224	210	210
R-squared	0.89	0.90	0.87	0.88
Test CRS (P-Values)		0.69		0.62

Robust standard errors reported below each coefficient. Within R-squared reported
(*) significant at 10%; (**) significant at 5%; (***) significant at 1%

The results (for both publications and citations) of estimating this model are presented in Table 5. The first column shows the estimated parameters without spillover effects; these results are statistically equivalent to those in Table 3. When we include the variable S$_{it}$, which aims to capture the spillover effect, the magnitude of the long-run elasticity for HERD remains stable (dropping only marginally). Interestingly, the variable S$_{it}$ is highly significant and has a large estimated parameter. In addition, the value of the time trend variable (Year) drops and is no longer significant. This validates our conjecture that the time trend was somehow capturing part of the spillover effect. In the case of citations (a quality adjusted measure of output) we obtain a positive and significant estimation for spillovers. It is interesting to note that in this case the time trend is negative and significant. This result can be interpreted as indicating an overall negative trend in the production of science output once

[13] We do not have enough observations to search for a different lag structure for the spillover variable.

adjustment has been made for impact. So, if we consider citations as a proxy for 'quality' of science and not just impact, the model indicates an overall decrease in the 'quality' of the scientific output at the world level.

Another interesting result relates to the magnitude of the coefficients for both domestic HERD and international spillovers (S_{it}). Their sum is very close to that for publications and just above that for the citation estimates. These results suggest the presence of decreasing returns to scale at the domestic level,[14] and constant or perhaps even increasing returns to scale at the global level. However, the null of constant returns to scale at world level was never rejected.

4.1 Testing for the contribution of proximity

The results for international spillovers obtained above rest on two assumptions: first that there are spillovers and, second, that the transmission mechanism is proximity. In our case the latter is measured on the basis of co-authorships between countries. In this section we question to what extent proximity is a real transmission mechanism. It is quite clear that because we are dealing here with public science, spillovers are expected 'almost by definition', to the extent that scientists read what other scientists around the world have published. This is the 'codified' component of the knowledge dissemination process. However, we know that not all knowledge created during a research programme is 'codified', an important proportion of it is assumed to be tacit. It is here that proximity or personal interactions among scientists become relevant. Hence, what follows can be interpreted as a test of the relative importance of codified versus tacit knowledge in science production.

For each country in the sample we replaced each component of the weighting matrix by a random draw from a uniform probability function. In this case, we use the uniform [0,1] density distribution, with the simulated coefficients standardised in order to add 1. After this we constructed the new stock of knowledge and ran the models as in the previous section. In order to guarantee independence of our results from each random draw, we reproduced the process 1,000 times.

If proximity matters for knowledge spillovers in science production, we should observe that the economic importance and statistical significance of the spillover coefficients obtained in our previous results are higher than those from the random matrix model. That is, we would expect that the majority of the spillover coefficients generated during our simulations would be 'to the left' of the coefficient estimated using our proximity matrix.

The results for publications are depicted in Figure 5 which presents the empirical distribution of the simulated spillover coefficients and compares them with the calculated value in Table 5 (represented by the perpendicular line on the graph). As can be seen from Figure 5, there is no evidence that proximity matters in the case of publications. Indeed, with 60% of the runs generating coefficients higher than the

[14] The result of decreasing to scale at domestic level could have interesting implications for the analysis of long-term economic growth. This result is consistent with the work of Jones (1995) showing that there is no correlation between R&D investment and total factor productivity (TFP) growth in the long run (although there is in the short run).

calculated ones, the opposite would seem to be the case. The results for citations are shown in Figure 6 and mirror those for publications.

Figure 5 Simulated spillovers, uniform random matrix, publications

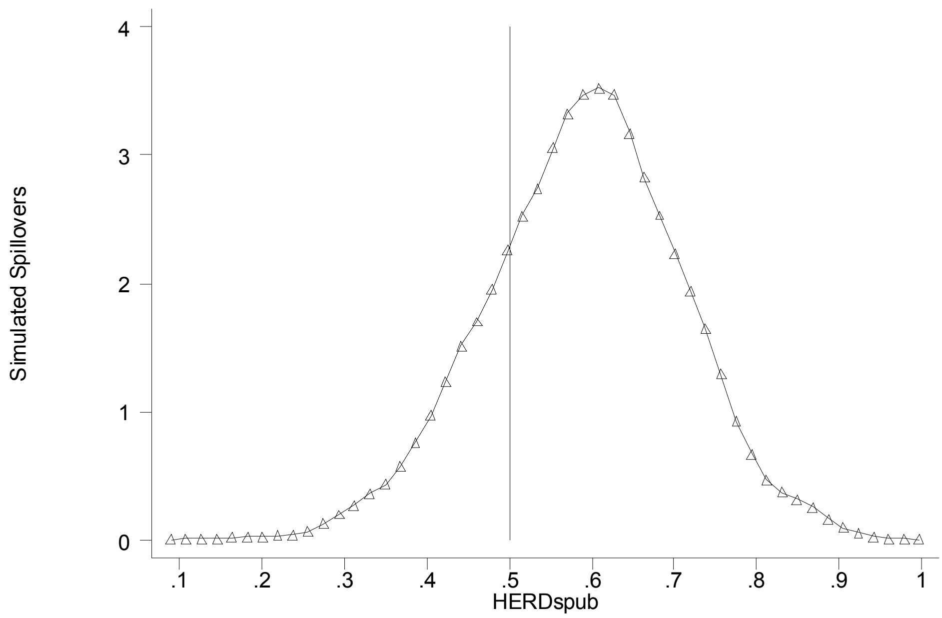
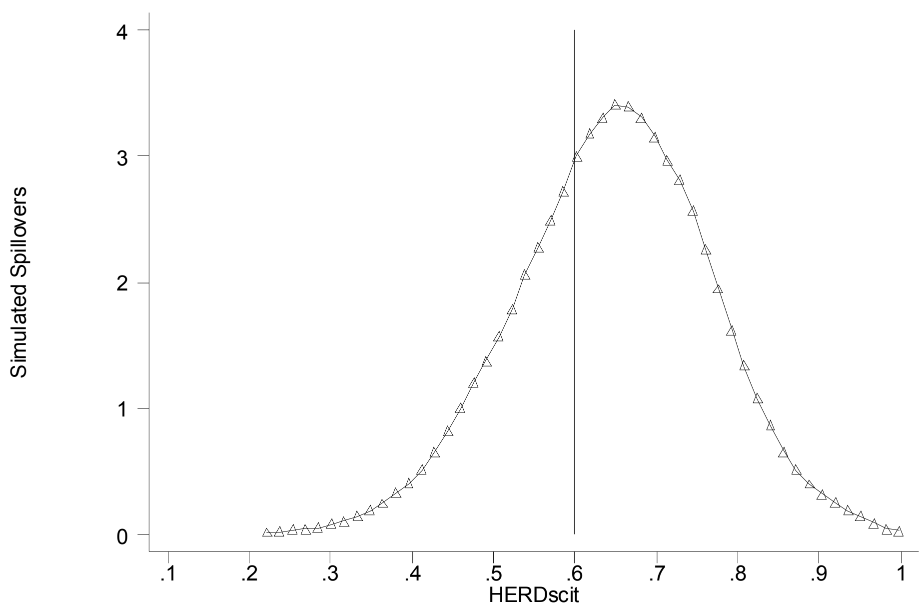


Figure 6 Simulated spillovers, uniform random matrix, citations



In summary, our results strongly reject the idea that proximity is important as a mechanism for transmitting spillovers in publicly funded research. According to our results, it seems that the 'codified' component of the knowledge clearly dominates. Of course, it is possible that the measure of proximity we used is not the correct one, in which case our test becomes one of the 'correctness' of using international co-authorship as an index of proximity, and alternative indices of proximity should be developed. This is part of our future research agenda.

4.2 The importance of the USA

The results in Table 4 indicate that there is a great propensity for scientists from several countries to publish with USA co-authors. In this section we explore more deeply the importance of the USA as an independent source of spillovers. We want to determine the extent of the dependence on the R&D investments made by the country that is the supposed world leader in several scientific disciplines. If this dependence is strong, the policy decisions of the rest of the world can no longer be taken in isolation. In other words, changes in US science policy (e.g. discontinuation of stem cell research, or surges in investment in military-related science) will affect the scientific production of the rest of the world. If the countries in the rest of the world want to continue with research that is not in line with reductions/changes in the USA's research budget, they will have to make extra efforts to compensate for the USA's knowledge decline (and therefore spillovers) in these particular fields.

In order to investigate these issues we re-estimated the spillover effect by splitting world level investment into two groups: USA and the Rest of the World (the rest of the countries being examined). We estimated the following knowledge production function:

$$y_{it} = \alpha_i + \beta_0 W(r)_{it} + \beta_1 W(r)_{USA,t} + \beta_2 W(r)_{-i-USA,t} + \gamma\ X_{it} + u_{it},\quad i = 1,.....N \qquad (7)$$

where y_{it} is the (log) output of the research 'intermediate' output (papers and citations) by country i (14 countries) and time t (21 years). $W(r)_{it}$ is (the log of) a distributed lagged function of real past R&D expenditure and X_{it} is a vector of the control variables. The two new variables, $W(r)_{USAt}$ and $W(r)_{-i-USA,t,}$ capture respectively the impact of the USA stock of knowledge in the research outputs of each country, and the impact of the stock of knowledge from all the remaining countries, except the USA and country i. In building these two new knowledge stocks we assumed the same lag structure and weights as before. The results for the Rest of the World were weighted assuming an equal contribution from each country to the non-USA pool of knowledge (given that our proximity/personal interaction measure was non significant).

As can be seen from Table 6, the source of the spillovers is important. For both publications and citations, spillovers from the USA are numerically more important than spillovers from the other countries. However, we also tested whether the differences between the spillover coefficients were statistically significant. The differences were not statistically significant for publications, while in the case of citations they were, but only marginally so. This result suggests that it is only in terms of quality/impact adjusted research output that spillovers from the US play a dominant role. Another interesting result is related to the impact on the returns to scale. The assumption of constant returns to scale is rejected both for publication and citations, pointing to the presence of increasing returns at the aggregate level.

Table 6: Results using 6 (7) lags of RD for Publications (Citations) plus Spillovers according to source of Origin [dependent variable log Publications (Citations)]. Fixed Effects by Country Included

	Publications	Citations
Non-$HERD_{it}$	-0.003	-0.01
	1.51	4.56***
Year	-0.012	-0.021
	1.86*	2.38**
$HERD_{it}$	0.418	0.468
	10.31***	11.27***
$HERD_{USA,t}$	0.494	0.585
	7.21***	8.08***
$HERD_{-i-USA,t}$	0.356	0.206
	2.20**	1.22
Constant	24.536	39.166
	2.00**	2.46**
Observations	224	210
Number of Country Code	14	14
R-squared	0.92	0.91
Test CRS (P-Values)	0.07*	0.09*
Test $HERD_{USA,t}$= $HERD_{-i-USA,t}$	0.49	0.07*

Robust standard errors reported below each coefficient. Within R-squared reported.
(*) significant at 10%; (**) significant at 5%; (***) significant at 1%

5 Conclusions

Modelling and measuring scientific production is not an easy task given the fact that science inputs and outputs are difficult to quantify in terms of both quantity and quality. Is a citation a realistic impact or quality adjusted proxy for the quality adjusted outputs of scientific research? How can we model the cumulative process of knowledge creation? Can it be assumed that some form of maximising behaviour for researchers will allow one to use a production function specification to examine scientific production?

This paper does not provide the final answers to these questions, but it acknowledges their relevance and tries to take into account their implications. The approach in this paper is pragmatic. It employs quantitative methods to examine variables believed to influence scientific production. Specifically, it develops econometric models based on the production function metaphor to relate a sub-set of inputs to two of the most common university research outputs: publications (as a proxy for the production of codified research) and citations (as an impact adjusted proxy for codified research production). This paper does not provide accurate or robust estimates of the elasticities of these outputs with respect to investment, a doubtful task given the poor quality of the data sources and the modelling problems. Instead, its aim is to explore two key features of the process of scientific production, the lag structure of output and the nature of knowledge spillovers.

To address the lag structure we estimated a polynomial distributed lag model. After determining the most appropriate specification for this model, we focused on the assessment of knowledge spillovers among the 14 countries considered. We found some evidence of a strongly positive long-run relationship between Higher education R&D (HERD) and the two research outputs examined. For both publications and citations we found evidence of decreasing returns to the domestic component of R&D. It should be noted that the parameters of the knowledge production function were very stable and robust to different compositions of the countries in the sample.

These results refer to long-run impacts. However, there is also a long and quite complex lag structure with regard to the impact of domestic R&D on the different research outputs. These weights are quite different to those typically assumed in standard econometric models. There is little evidence of any positive impact in under two years in the case of publications and three years in the case of citations. The total cumulated effects (the long-run elasticities) are spread over 6 and 7 years respectively, and reach a maximum towards years 5 and 6 respectively. Evaluation of the impact of science policies on research outputs has to take account of this lag structure if erroneous conclusions are to be avoided.

The analysis of international spillovers indicates that for publications and citations there is evidence of a significant impact from the weighted investment in HERD in other countries. We studied the mechanism of transmission of international spillovers using two different approaches. One approach, based on an index of scientific proximity (personal interaction) between countries was rejected. It seems that the 'codified' component of the knowledge is clearly dominant. However, it is possible that our measure of proximity (based on international co-authorship) is not the correct one. A second transmission mechanism was investigated by splitting the world pool of knowledge between the USA and the rest of countries considered. In this case the results were consistent with the hypothesis that spillovers from the USA (the scientific leader in many disciplines) are higher than those from the other countries considered. This result, however, is only significant in the case of citations. For this model, the assumption of constant returns to scale is rejected both for publications and for citations, pointing to the presence of increasing returns at the aggregate level (by contrast with the country level results).

From a policy perspective our results highlight the importance of two phenomena. First, the lag between investment and the research output is considerable, which is very problematic because it implies that there is a set of factors for which we could not control that may have a changing impact on the research outputs. Consequently, if it becomes difficult to link the inputs and outputs of the scientific production 'box', connecting socio-economic benefits to certain scientific investments becomes a doubtful enterprise. If quantitative approaches to the assessment of public investment in science are applicable, the most promising area of development is one that would focus on a better understanding of the relationship between the inputs and the most direct outputs of the scientific process.

This work provides some quantitative evidence supporting the view that science is an international enterprise characterised by major spillovers across countries. Tackling the funding, management and organisation of science from only a national viewpoint will lead to incomplete or erroneous results. The evidence of the significant and extremely important spillovers from the US science system underscore the hazard of other OECD countries defining their science priorities without taking cognisance of changes in priorities in the US system. A reduction in public spending on stem cell research or an increase in spending on military research such as those that have occurred in the US in recent years affects in a significant way not only the US science output, but also those of other countries, involving significant changes in public spending by these other countries to avoid a reorientation of their science output.

The results of our model estimations show that (excluding Canada), although there are knowledge spillovers across European countries, the impact of these is lower (at least in terms of impact/quality adjusted measurement) than the impact of US spillovers. This may be due to the size of the science investment in the US (142 € per capita against 89 € per capita in the EU-15 in 1999), but also to the fact that although the EU has a similar or even higher publication output than the US, the EU countries achieve excellence only in a small number of fields (EC, 2003). These results seem to support the policy view of the need for a European Research Council able to fund (at appropriate levels) research excellence at the European level.

References

ADAMS, J. (1990): 'Fundamental Stocks of Knowledge and Productivity Growth,' *The Journal of Political Economy*, 98, 673-702.

ADAMS, J., AND Z. GRILICHES (1998): 'Research Productivity in a System of Universities', *Annales d'Economie et de Statistique*, V0, 49-50,(Jan-Jun 1988), 127-162.

ARUNDEL, A., AND A. GEUNA (2004): 'Proximity and the Use of Public Science by Innovative European Firms', *Economics of Innovation and New Technology*, 13, 559-580.

BONACCORSI, A., AND C. DARAIO (2003): 'A Robust Nonparametric Approach to the Analysis of Scientific Productivity.' *Research Evaluation*, 12, 47-69.

CRESPI, G., AND A. GEUNA. (2004): *The Productivity of Science*. Brighton: SPRU Report prepared for the Office of Science and Technology (OST), Department of Trade and Industry (DTI), UK. http://www.sussex.ac.uk/spru/publications/reports/CrespiOST.pdf

DASGUPTA, P., AND P.A. DAVID (1994): 'Towards a New Economics of Science', *Research Policy*, 23, 487-521.

EUROPEAN COMMISSION (2003): *Third European Report on Science & Technology Indicators*, Belgium.

FELIPE, J., AND F.M. FISHER (2003): 'Aggregation in Production Functions: What Applied Economists Should Know', *Metroeconomica*, 54, 208-262.

FISHER, F.M. (1969): 'The Existence of Aggregate Production Functions', *Econometrica*, 37, 553-576.

GEUNA, A. (1999): *The Economics of Knowledge Production: Funding and the Structure of University Research*, Cheltenham: Edward Elgar.

GEUNA, A. (2001): 'The Changing Rationale for European University Research Funding: Are there Negative Unintended Consequences', *Journal of Economic Issues*, 35, 607-632.

GEUNA, A., AND L. NESTA. (2005): 'University Patenting and its Effects on Academic Research: the Emerging European Evidence', *Research Policy*, Forthcoming http://www.sussex.ac.uk/Users/prff0/Publications/Geuna_Nesta%20University%20patenting.pdf

GREENE, W. (1993): *Econometric Analysis*. Englewood Cliffs, NJ: Prentice Hall.

GRILICHES, Z. (1979): 'Issues in Assessing the Contribution of Research and Development to Productivity Growth', *The Bell Journal of Economics*, 10, 92-116.

GUELLEC, D., AND B. VAN POTTELSBERGHE (2001): ‘R&D and Productivity Growth: Panel Analysis of 16 OECD Countries’, *OECD Economic Studies*, 0(33), 103-126.

HALL, B., A. JAFFE, AND M. TRAJTENBERG (2001): ‘The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools’, *NBER Working Paper Series*, 8498.

JACOBSSON, S., AND A. RICKNE (2004): ‘How Large is the Swedish 'Academic' Sector Really? A Critical Analysis of the Use of Science and Technology Indicators,’ *Research Policy*, 33, 1355-1372.

JAFFE, A. (1989): ‘Real Effects of Academic Research’, *American Economic Review*, 79, 957-970.

JOHNES, J., AND G. JOHNES (1995): ‘Research Funding and Performance in UK University Departments of Economics: A Frontier Analysis’, *Economics of Education Review*, 14, 301-314.

JONES, C. (1995): ‘R&D-Based Models of Endogenous Growth’, *Journal of Political Economy*, 103, 759-784.

KOLENIKOV, S. (2000): ‘Icomp: Information Complexity Measures’, http://ideas.repec.org/c/boc/bocode/s410201.html, accessed on 02/06/2004.

KLETTE, T., AND F. JOHANSEN (2000): ‘Accumulation of R&D Capital and Dynamic Firm Performance: a Not-so-Fixed Effect Model', in D. Encaoua, B. Hall, F. Lisney AND J. Mairesse (eds): *The Economics and Econometrics of Innovation*, Kluwer Academic Publishers, Boston/Dordrecht/London, 367-397.

KLEVORICK, A., A. LEVIN, R. NELSON, AND S. WINTER (1995): ‘On the Sources and the Significance of Interindustry Differences in Technological Opportunities’, *Research Policy*, 24, 185-205.

MANSFIELD, E. (1991): ‘Academic Research and Industrial Innovation’, *Research Policy*, 20, 1-12.

MOWERY, D., R. NELSON, B. SAMPAT, AND A. ZIEDONIS (2004): *Ivory Tower and Industrial Innovation*, Stanford: Stanford Business Books.

NARIN, F., K. HAMILTON, AND D. OLIVASTRO (1997): ‘The Increasing Linkage between U.S. Technology and Public Science,’ *Research Policy*, 26, 317-330.

NSF (2002): ‘Science & Engineering Indicators’.

OECD (2003a): ‘*Turning Science into Business. Patenting and Licensing at Public Research Organisations*’, Paris: OECD.

OECD (2003b): '*The Sources of Economic Growth in OECD Countries*’, Paris: OECD.

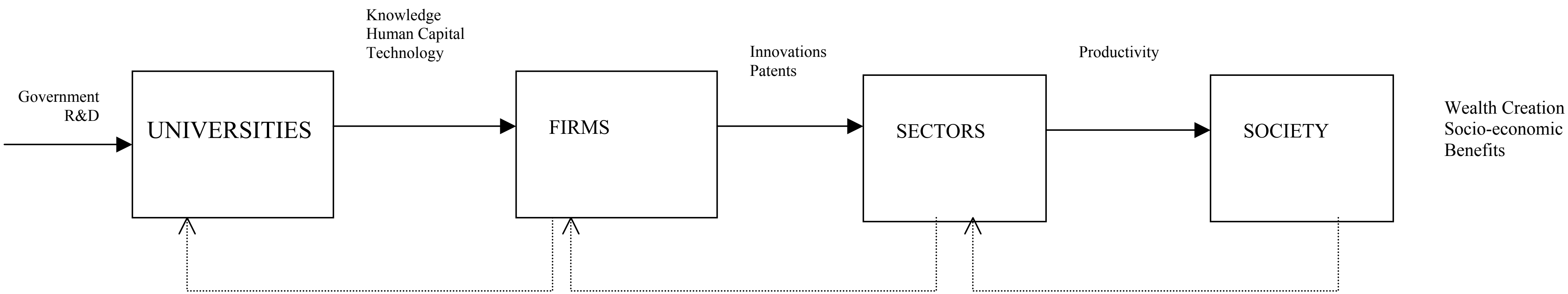
OECD (2002a): '*Benchmarking Industry-Science Relationships*', Paris: OECD.

OECD (2002b): '*Frascati Manual: Proposed Standard Practice for Surveys on Research & Development Expenditures*', Paris: OECD, 254.

SCOTT, A., G. STEYN, A. GEUNA, S. BRUSONI, AND W.E. STEINMUELLER (2001): '*The Economic Returns to Basic Research and Benefits of University-Industry Relationships*'. A report commissioned by the OST-DTI, SPRU, University of Sussex, Brighton, 2001. http://www.sussex.ac.uk/spru/publications/econreturnsost.pdf.

STEINMUELLER, W.E. (2001): 'Economics of Science', in *International Encyclopaedia of Social and Behavioral Sciences*, Vol. 20; Oxford: Elsevier, pp.13664-13668.

Figure 1


Government R&D
UNIVERSITIES
Knowledge
Human Capital
Technology
FIRMS
Innovations
Patents
SECTORS
Productivity
SOCIETY
Wealth Creation
Socio-economic
Benefits