PRELIMINARY AND INCOMPLETE - PLEASE DO NOT QUOTE

Agglomeration and network effects in regional R&D productivity

Attila Varga
Department of Economics and Regional Studies
Faculty of Business and Economics
University of Pécs, Pécs, Hungary

Dimitrios Pontikakis
Joint Research Centre – Institute of Prospective Technological Studies
European Commission
Seville, Spain

Georgios Chorafakis
Research Directorate-General
European Commission
Brussels, Belgium
and
Department of Geography,
University of Cambridge, Cambridge, United Kingdom

Abstract

This paper reflects on the increasing recent interest in the role that (geographical and sectoral) specialisation could possibly play in the future allocation of R&D resources in Europe. So far the debate is more at the level of principles than based on facts. This paper contributes to the debate by following an empirical approach. It further develops the methodology applied in Varga (2000) to set up a spatial econometric modelling framework for estimating the dynamic effects of agglomeration and interregional scientific networks on R&D productivity in regional knowledge creation (measured by patent applications and publications) at the level of EU NUTS-2 regions. This empirical modelling framework is applied to delineate the "critical mass" of agglomeration in Europe as well as to classify EU regions into different tiers according to the strengths of their agglomeration effects. These effects are then compared to the network effects of interregional connectedness as promoted by the EU's framework programme for research (FP). The estimated model is used then for an assessment of the impacts of EU Framework Program expenditures on technological development and for carrying out policy impact simulations.

1. Introduction

A very recent debate in EU policy cycles is concerned with the optimal spatial and thematic allocation of resources for research (Georghiou, 2008; Martin, 2008; Foray, 2008; Cooke, 2008; Pontikakis, Chorafakis and Kyriakou, 2008). This stems from a concern that EU research funds are spread too thinly across Europe without achieving the impact on growth and employment that is expected of them. 'Smart specialisation' or the spatial and thematic concentration of R&D resources on the basis of existing patterns of technological specialisation, is therefore put forward as one possible solution (Foray and Van Ark, 2007).

Localised interactions play a central role in innovation (Jaffe, Trajtenberg, Henderson 1993; Anselin, Varga, Acs 1997). The importance of scale effects on the productivity of regional innovation systems has also been acknowledged in the literature. Varga (2000) identifies a threshold level of local agglomeration, above which increases in the marginal productivity of R&D inputs can be expected. This suggests that regional research efforts can have a much more meaningful impact where a critical mass of agglomeration (in terms of business services and high-technology employment) are also present.

On the other hand however, other studies emphasise the role of networks and evolving structural features and highlight the systemic nature of regional knowledge production processes (Maggioni, Nosvelli and Uberti 2007; Ponds, Oort and Frenken 2008). Importantly, different types of research impose different requirements on scale and place a different emphasis on tacit knowledge and by extension, proximity (Malmberg and Maskell, 1997). These bring up the possibility that, contingent on sufficient linkages to international knowledge production networks, regions may be productive even in the absence of local agglomeration. If this is the case, then increasing interregional connectedness may be an appropriate pathway to improving R&D productivity.

In this paper we aim to explore the effects of regional agglomeration and interregional networking on the production of technological and scientific knowledge across EU regions. To this purpose we conduct an estimation of an enhanced Romer-Jones knowledge production function at the EU regional level, within a framework that also takes into account possible spatial dependence effects, with a view to estimate how R&D productivity varies across regions according to their local innovation capacity. A discussion of our findings then follows. As this is work in progress, the paper concludes with our research plans for the immediate future.

2. Methodology: an empirical framework

Given that we want to relate our research to the ongoing debate of whether geographic specialisation of research expenditures is a reasonable reaction to current problems in Europe it is important to try to describe the dynamic effects of R&D support as closely as data allow. To this end, we employ a multiple equation system[1] encompassing a knowledge production function (KPF), a function modelling the parameter (beta) of the R&D variable in the KPF, a function estimating the spatial location of R&D expenditures, a function estimating the spatial location local innovation capacity.

Our starting point is the basic KPF initially specified by Romer (1990) and parameterised by Jones (1995):

$$\text{Eq.1} \quad dA = \delta\, H_A^{\lambda}\, A^{\varphi},$$

Where, H_A refers to research inputs (e.g. number of researchers or research expenditures), A refers to the total stock of technological knowledge (codified knowledge component of knowledge production in books, patent documents etc.), dA refers to the change in technological knowledge, δ is the "research productivity parameter" ($0<\delta<1$), φ is "***codified knowledge spillovers*** parameter" (reflects spillovers with unlimited spatial accessibility) and λ is the "***research spillovers*** parameter" (reflects the geography of knowledge spillovers).

Eq. 1 suggests that an increase in the resources devoted to research (H_A) or to the amount of existing knowledge accessible to researchers (A) will have a positive effect on technological change (dA). We adopt here the spatial interpretation of the Romer equation outlined in Varga (2006: 1175) which, drawing from literature on regional and urban economics and the new economic geography, assumes that the parameter of H_A, λ ***changes with the geographic concentration of economic activities.*** The size of the parameter could vary, depending on the balance between positive and negative agglomeration economies.

Eq. 2 describes regional knowledge production, specifying K_r, the amount of new knowledge produced in region r as some function of R&D inputs and other factors Z_r such as the level of regional agglomeration[2].

$$\text{Eq.2} \quad K_r = K\,(RD_r, Z_r)$$

As emphasised by as diverse strands of literature as endogenous growth accounting, innovation systems and the new economic geography, knowledge production is a dynamic, ***cumulative process***. This process is described by Eqs. 3 and 4a,b. In a static context (Eq. 3), research already in the region (RD) and agglomeration (Z_r) affect R&D productivity across regions ($\partial K_r/\partial RD_r$).

[1] This framework draws heavily from Varga (2006), where the rationale for the specification of equations is explained in further detail.
[2] Alternatively Z may reflect network effects: its precise specification is an empirical matter.

$$\text{Eq.3} \quad \partial K_r/\partial RD_r = f(RD_r, Z_r)$$

However, changes in RD_r and Z_r (dRD_r and dZ_r) are themselves determined by manifested R&D productivity and research already in the region respectively. This dynamic effect is described in equations 4a and 4b: in Eq. 4a the geographic distribution of R&D expenditures is a function of R&D productivity across different regions whereas in Eq. 4b the level of regional agglomeration is a function of research already in the region.

$$\text{Eq.4a} \quad dRD_r = R(\partial K_r/\partial RD_r)$$
$$\text{Eq.4b} \quad dZ_r = Z(RD_r)$$

At the national level, the magnitude of λ, reflecting the impact that the same number of researchers has on technological change, depends on the geographic structure of H_A (where $H_A = \Sigma_r RD$).

$$\text{Eq.5} \quad \lambda = \lambda\,(GSTR(H_A))$$

Correspondingly, the rate of national technological change may be described in the Romer-Jones fashion as:

$$\text{Eq.6} \quad dA = \delta\, H_A{}^{\lambda} A^{\varphi}$$

The above equations may be tested empirically, subject to the availability of appropriate data. We proxy R&D inputs with data on regional R&D expenditures and new knowledge flows with data on patent applications to the EPO with at least one inventor based in the region.

In order to test empirically these relationships we use the following econometric specifications. In terms of functional form we opt for a (natural) logarithmic conversion of all our variables, which has the appealing quality of allowing the interpretation of coefficients as elasticities[3]. Using subscripts *i* and *N* to denote individual regions and nations (in our case EU member states) respectively, the regional KPF may be specified as:

[3] This functional form is common in empirical specifications of Romer-type KPFs (see Porter and Stern, 2000; Furman et al., 2002; Varsakelis, 2006). Taking logarithms also has the added advantage of minimising the influence of outliers and normalising measures expressed in different units of measurement.

Eq. 7. $Log(K_i) = \alpha_0 + \alpha_1 Log(RD_i) + \alpha_2 Log(PATSTOCK_N) + \varepsilon_i$,

Eq. 8. $\alpha_{1,i} = \beta_0 + \beta_1 Log(\delta_i) + \beta_2 Log\ (NET_i)$

Eq. 9 $Log(K_i) = \alpha_0 + \beta_0 Log(RD_i) + \beta_1 Log(\delta_i) Log(RD_i) + \beta_2 Log(\delta) Log(RD_i)$
$\alpha_2 Log(PATSTOCK_N) + \varepsilon_i$,

where

$\alpha_{1,i}$ is regional research productivity

δ_i is regional innovation capacity

NET_i is interregional research networks

$PATSTOCK_N$ is national patent stock

The above are essentially empirical tests of equations 2 and 3.

As most measures of absolute concentration of economic activity introduce multicollinearity, they are likely to be problematic in a regression context with interaction terms. We overcome this problem by using a novel index of regional innovation capacity.

Regional innovation capacity δ is proxied by a novel index of agglomeration of knowledge intensive employment. Our index is a size-adjusted (in the spirit of the index developed by Elison and Glaeser (1997)) variation of the popular location quotient (LQ) measure and is calculated as:

$$\delta_i = [(EMPKI_i / EMPKI_{EU}) / (EMP_i / EMP_{EU})] / [(1 - \sum_j (EMPKI_{i,j} / EMPKI_{j,EU})][1 - (EMP_i / EMP_{EU})]$$

Where EMPK is employment in knowledge intensive economic sectors[4], EMP is total employment and the subscripts *i* and *EU* stands for region and EU aggregate respectively. Just like the LQ, δ has the interesting property of taking a value of 1 for regions with a level of agglomeration close to the EU average. However, unlike the LQ, in δ the denominator is designed in such a way as to penalise small regions, by yielding higher values for regions with a higher level of employment. As δ captures economic activity that is heavily involved not only in the production but also in the diffusion, assimilation and productive deployment of knowledge, we consider it an appropriate indicator of local innovation capacity.

Following on, the determinants of the location of R&D expenditures (RD_i) and regional innovation capacity (δ_i) may be empirically tested by:

[4] The classification of knowledge intensive economic sectors (devised by Eurostat) includes: high and medium high technology manufacturing (NACE 1.1 sectors xxxxx), high technology services, knowledge intensive market services (NACE 1.1 sectors 61, 62, 70, 71, 74), financial services (NACE 1.1 sectors 65, 66, 67), amenity services – health, education, recreation (NACE 1.1 sectors 80, 85, 92)

$$\text{Eq. 10} \quad Log(RD_i) = \lambda_0 + \gamma_1 Log(\alpha_{1,i}) + \lambda_1 Log(PATSTOCK_{N,i}) + u_i$$

$$\text{Eq. 11} \quad Log(\delta_i) = \xi_0 + \xi_1 Log(RD_i) + \xi_2 Log(PATSTOCK_{N,i}) + \mu_i$$

This framework allows for testing various alternative hypotheses.

First, by substituting agglomeration proxies for network proxies the same modelling framework can be used to compare the relative importance of agglomeration and network effects.

Second, bearing on mind the sharp differences in the worlds of scientific and technological research and borrowing the terminology used by Stokes (1997) and Mokyr (2002), we consider two distinct types of research and examine them separately:

(a) **Edison-type**: research the products of which have clear economic applications, pursuing market-oriented innovation. Sometimes dubbed "*competitive research*" among EU policy analysts.
(b) **Pasteur-type**: science-oriented research, mediated by the distinct rules and incentives of the modern scientific establishment. Sometimes dubbed "*pre-competitive research*" among EU policy analysts (and referred as such in relevant EU treaties).

Edison-type research frequently results in patents, while the findings of Pasteur-type research are commonly documented in scientific publications. We use patents and publications in separate KFPs to draw our comparisons.

As a further step, on the basis of the regression estimates, simulations for different policy scenarios will be used to explore the likely effects of FP funding across EU regions (to be explored in the near future).

3. Data and Estimation Issues

Our estimates are based on regional-level data for 190 European regions (a mixture of NUTS2 and NUTS1 regions) where information was complete enough for the purposes of our study (see Appendix 2 for a list of regions). Data on patenting, R&D expenditures, high-tech employment and other indicators of regional knowledge infrastructure are drawn from Eurostat´s New Cronos database. Data on scientific publications are from the European Commission's Regional Key Figures[5] (RKF) database and were originally compiled by CWTS, Leiden University. The study also makes use of original data, drawn from the Commission's internal FP databases on regional knowledge linkages to construct an indicator of interregional network

[5] RKF is a novel database combining publicly available science, technology and innovation indicators from various disparate sources with indicators that have been developed for Commission use. The database has been commissioned by DG RTD and is implemented by Fraunhofer ISI.

intensity and, using the knowledge production function, compare its effect to that of geographic specialisation. Using the FP5 database we have constructed an *n* by *n* matrix (where *n*=number of NUTS 2 regions in the sample) where a matrix element with a value 1 means a common FP project of two regions and zero otherwise. This matrix is used to calculate the average R&D expenditures of network partner regions for each region.

Although patent counts are far from a perfect proxy of innovation (Griliches, 1990), they are the only measure that is available for a large number of European regions and over a number of years. Comfortingly, previous research has shown that at the level of regions, patent counts correlate well with innovation counts (Acs, Anselin and Varga, 2002) and both measures provide very similar results in the knowledge production function context.

We use a cross-section of EU regions as opposed to a panel, due to lack of data in the time-series dimension. A quick glance over the regional R&D and patenting data for Europe indicates that, even where the data is complete, variation in the time-series dimension is small. Dependent variables are lagged by one period, reflecting the dynamic nature of our system. Temporally lagged dependent variables have the added advantage of partially countering potential endogeneity problems. Spatial econometric tests and, where appropriate, adjustments are made in all estimates.

Three different types of national patent stocks were constructed and tested empirically: patent stocks with no depreciation (Porter and Stern, 2000; Furman, Porter and Stern, 2002), and, using the perpetual inventory method (PIM), patent stocks with a 13 per cent (Park and Park, 2006) and a 15 per cent annual depreciation rate (Hall, 1993) respectively. Non-depreciated stocks are simply the cumulative number of patent applications from 1992 on, while PIM estimates of contemporary patent stocks are based on the following formula:

$$PATSTOCK_t = PATSTOCK_{t-1} * (1 - d) + PAT_t$$

Where PAT_t are contemporary patent flows and d is the depreciation rate (0.13 or 0.15). Like other independent variables all patent stock variables are lagged by one period. Initial stocks are set in year 1998 taking into account compound annual growth in the five preceding years[6]. Descriptive statistics are in Appendix 1.

[6] Initial stock equals flows for first year divided by the sum of compound growth for the preceding five year period and the depreciation rate. As all patent stock variables are time t-1, annual compound growth rates for the PIM variables were calculated for the 5 year period 1992-1997. Exceptions are Malta and Lithuania, where due to lack of data in the time series dimension, the preceding 4 year period (1993-1997) was used instead. For the non-depreciated stocks, a value of 1 was assumed in the case of Lithuania for 1992 (which is close to the average for that country in the following two years), while the 1998 value was estimated as the average of 1997 and 1999.

4. Preliminary Results

We begin with the estimation of the KPF using patents as a proxy of new knowledge produced in each European region in the year 2002 (Table 1). Regressions were estimated in Eviews and Spacestat. To begin with, regression diagnostics indicate no problems with multicollinearity, as the multicollinearity condition number for all models is below the rule-of-thumb threshold of 30.

The first baseline model (1) confirms that, on average, lagged gross regional R&D expenditures (GRD2001) have a significant relationship with contemporary patent flows. Moreover, the proximity of the estimated coefficient to unity suggests that innovation flows throughout European regions are on average about proportionate to R&D inputs. Model 2 includes the product of lagged R&D expenditures and δ, our index of regional innovation capacity. Model 2 suggests that local innovation capacity has a positive, statistically significant and quantitatively distinct effect on R&D productivity, confirming the significance of agglomeration effects. Interpreted from an innovation systems perspective, this finding reflects the importance of knowledge interactions between different institutional actors engaged in knowledge-intensive economic activities (e.g. users versus producers, academic institutions, government actors etc) for the innovation (Andersen, 1992; Nelson, 1993; Edquist, 1997; Cooke, 2001). The importance of colocation is also suggestive of the significance of tacit knowledge (Malmberg and Maskell, 1997). Taken together with the fact that the inventive performance of adjacent regions does not seem to affect a region's innovation flows this seems to be consistent with the idea of national boundaries as a break to knowledge spillovers, though of course it is does not conclusively demonstrate its veracity.

Models 3 and 4 test the significance of network effects, by including the product of gross regional R&D expenditures times the average value of the R&D expenditures with which region *i* had at least one joint research project in FP5 (Log(GRD2001)*Log(NET)). The product term is statistically insignificant with (4) or without (3) the agglomeration product. This result suggests that the average R&D expenditures of collaborating regions do not affect R&D productivity in the region.

Model 5 introduces national patent stocks (PSTCK2001), indicating that historically accumulated technological knowledge has a positive, statistically significant and quantitatively distinct effect on regional patenting. Interestingly, the coefficient of Log(GRD2001)*Log(δ) drops from around 2.6 in models 2 and 4 to about 1.7, suggesting that codified knowledge spillovers capture at least some of the effect attributed to agglomeration in the previous models.

In models (1-5), tests for spatial dependence indicated no influence from adjacent proximity (Neighb), but some role for distance (INV1 and INV2). In model (6), controlling for spatial dependence, the substantive results remain, although the value of the coefficient for the agglomeration interaction term is smaller and its statistical significance drops to 95 per cent. Overall, the inventive performance of adjacent regions does not seem to affect a region's innovation flows. However, the importance of distance hints at some other important spatial limitation, possibly national boundaries. It is worth noting that all models explain 70 per cent or more of the variation in regional patenting, with model 4 exhibiting the highest adjusted adjusted R-squared (0.79). In the final model there is an indication of remaining spatial error dependence which will be the subject of further refinements in our research.

Table 1. Regression Results for Log (Patents) for EU regions, 2002 (N=190)

Model	(1)	(2)	(3)	(4)	(5)	(6)
Estimation	OLS	OLS	OLS	OLS	OLS	IV-Spatial Lag
Constant	-2.06***	-0.951***	-2.377***	-1.274**	-2.475***	-2.677***
	(0.316)	(0.416)	(0.431)	(0.502)	(0.445)	(0.429)
W_Log(PAT2002)						0.041***
						(0.010)
Log(GRD2001)	1.131***	0.941***	1.283***	1.096***	0.830***	0.837***
	(0.054)	(0.071)	(0.150)	(0.152)	(0.067)	(0.064)
Log(GRD2001)*Log(δ)		0.267***		0.268***	0.171***	0.139**
		(0.068)		(0.068)	(0.064)	(0.062)
Log(GRD2001)*Log(NET)			-0.0001	-0.0001		
			(0.0001)	(0.0001)		
Log(PSTCK2001)					0.238***	0.159***
					(0.037)	(0.040)
R^2-adj	0.70	0.72	0.70	0.72	0.77	0.79
Multicollinearity Condition Number	7	10	22	24	13	
Jarque-Bera test on normality of errors	3582***	5778***	3529***	5662***	12822***	
White test for heteroskedasticity	0.195	0.537	2.842	5.506	1.619	
LM-Err						
Neighb	4.165	2.591	3.259*	1.912	0.671	0.052
INV1	59.915***	46.40***	65.397***	40.966***	10.502***	1.445
INV2	62.669***	34.699***	58.886***	32.400***	12.896***	6.027**
LM-Lag						
Neighb	0.157	0.123	0.133	0.151	0.489	
INV1	52.502***	43.583***	51.876***	42.536***	18.388***	
INV2	67.106***	42.889***	64.697***	40.952***	10.439***	

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; W_Log(PAT2002) is the spatially lagged dependent variable where W stands for the weights matrix INV1; instruments in the IV-Spatial Lag estimation are W_Log(GRD2001), W_Log(δ), W_[Log(GRD2001)*Log(δ)] and W_[Log(GRD2001)*Log(NET)], where W stands for the weights matrix INV1. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

Table 2 estimates the KPF with scientific publications for year 2002 as the dependent variable. In all models, regression diagnostics indicate no problems with multicollinearity and, as with patents, the KPFs explain more than 70 per cent of variation in the data. Gross regional R&D expenditures explain most of the variation, with a coefficient in model 1 (0.94) suggestive of almost constant returns to scale. Strikingly, agglomeration effects appear to have no statistically significant influence on scientific R&D productivity (Models 2 and 4), while network effects (Models 3, 4 and 5) exert a statistically significant and quantitatively distinct influence on scientific R&D productivity.

In the case of Pasteur-type research at least, interregional networking is more important than local agglomeration. In other words, regions can perform research efficiently even in the absence of local agglomeration. Patent stocks are insignificant (Model 5), though ideally one would want to test the influence of publication stocks in this setting (something we plan to do in the near future). The fact that none of the spatial dependence measures is statistically significant, confirms the importance of codified (as opposed to tacit) knowledge for scientific research. The remaining heteroscedasticity will need further attention.

Table 2. Regression Results for Log (Publications) for EU regions, 2002 (N=189)

Model	(1)	(2)	(3)	(4)	(5)
Estimation	OLS	OLS	OLS	OLS	OLS
Constant	1.352***	1.113***	2.408***	2.167***	2.378***
	(0.231)	(0.316)	(0.296)	(0.356)	(0.308)
Log(GRD2001)	0.943***	0.984***	0.440***	0.482***	0.422***
	(0.039)	(0.054)	(0.103)	(0.108)	(0.114)
Log(GRD2001)*Log(δ)		-0.058		-0.061	
		(0.052)		(0.049)	
Log(GRD2001)*Log(NET)			0.0004***	0.0004***	0.0004***
			(7.18E-05)	(7.17E-05)	(7.47E-05)
Log(PSTCK2001)					0.011
					(0.030)
R^2-adj	0.75	0.75	0.78	0.78	0.78
Multicollinearity Condition Number	7	10	22	24	28
Jarque-Bera test on normality of errors	35***	44***	37***	45***	37***
White test for heteroskedasticity	18.054***	26.661***	25.774***	19.688***	33.071***
LM-Err					
Neighb	1.681	1.354	2.890	2.599	2.556
INV1	0.612	0.315	0.111	0.041	0.174
INV2	0.545	0.628	0.009	0.006	0.005
LM-Lag					
Neighb	2.777*	2.602	2.056	1.276	2.258
INV1	6.028**	5.40**	2.636	2.062	3.196
INV2	1.421	1.029	0.780	0.449	0.781

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

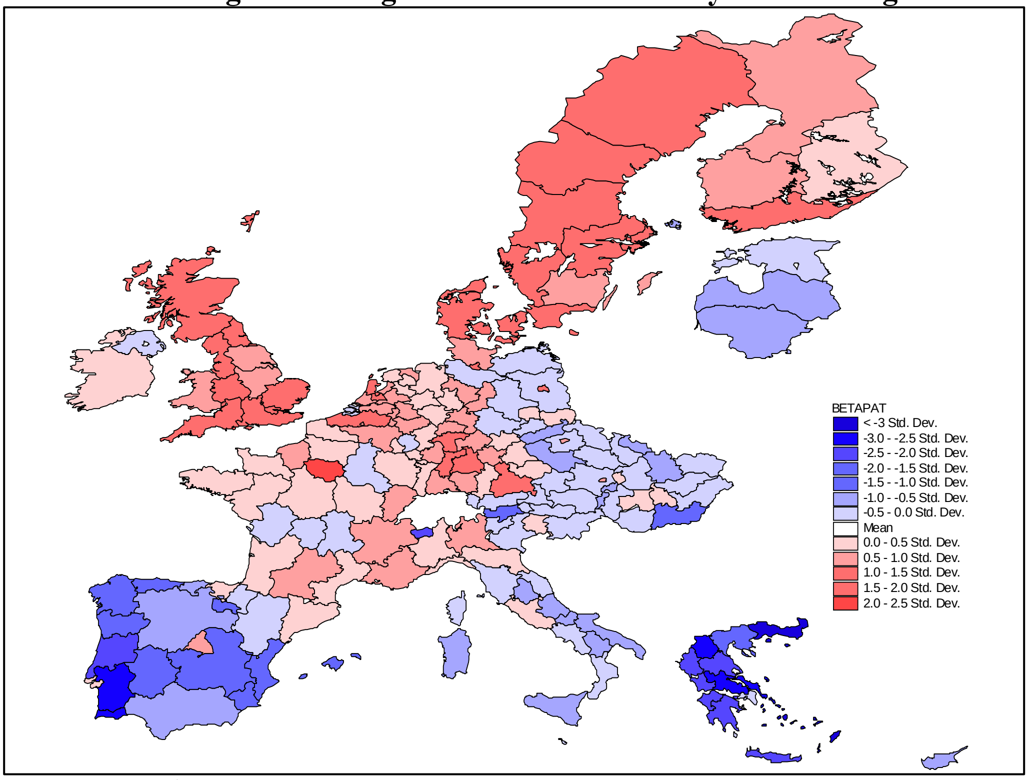
On the basis of the above models, we estimated regional productivity of research in innovation and scientific output using the following formulae:

$$BETAPAT_i = 0.837 + 0.139*\log(\delta_i)$$

$$BETAPUB_i = 0.422 + 0.0004*\mathrm{Log}(NET_i)$$

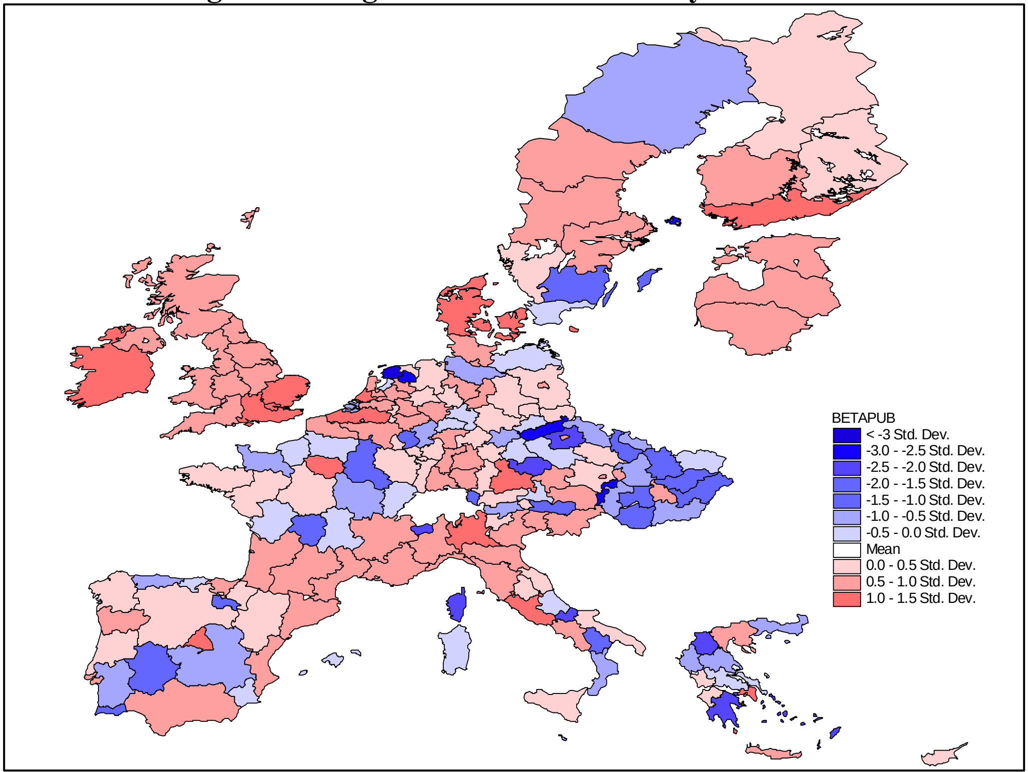
Our estimates are depicted in the following two maps, expressed in standard deviations from the European mean (Figure 1 and 2).

Figure 1 – Regional R&D Productivity in Patenting



Source: own estimates

Figure 2 – Regional R&D Productivity in Publications



Source: own estimates

We now move on to test the effect of R&D productivity on R&D expenditures. Regressions in Table 3 confirm that the spatial allocation of R&D expenditures is conditioned by R&D productivity, both technological (BETAPAT) and scientific (BETAPUB). Patent stock enters in models 3 and 4 It is noteworthy, spatial dependence is not an issue in any of the models in Table 3, suggesting that the relationship is localised within the boundaries of the region.

Table 3. Regression Results for Log (GRD2002) for EU regions, 2002 (N=190)

Model	(1)	(2)	(3)	(4)
Estimation	OLS	OLS	OLS	Heteroscedastic Error
Constant	9.576***	10.224***	8.379***	8.241***
	(0.284)	(0.221)	(0.429)	(0.406)
Log(BETAPAT)	20.089***	13.504***	10.155***	9.694***
	(1.404)	(1.198)	(1.319)	(1.242)
Log(BETAPUB)		5.876***	6.074***	6.062***
		(0.494)	(0.468)	(0.440)
Log(PSTCK2001)			0.140***	0.143***
			(0.029)	(0.027)
R^2-adj	0.52	0.73	0.76	0.76
Multicollinearity Condition Number	7	9	19	
Jarque-Bera test on normality of errors	54***	57***	58***	
White test for heteroskedasticity	1.001	9.326*	28.589***	
Wald test on heteroscedasticity				0.424
LM-Err				
Neighb	0.136	0.201	0.026	0.002
INV1	5.998	1.660	0.029	0.023
INV2	11.273	2.176	0.155	0.176
LM-Lag				
Neighb	0.231	0.083	0.058	0.056
INV1	0.002	5.517	0.220	0.527
INV2	0.012	3.089	0.222	0.175

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

In Table 4, we present our estimates of R&D and national technological knowledge stock on regional innovation capacity (δ). Both variables exert a positive and statistically significant effect on regional innovation capacity. Spatial dependence (INV and INV2) appears to be an issue with models 1 and 2. The remaining heteroscedasticity in the spatial error model will need further refinements in the estimations.

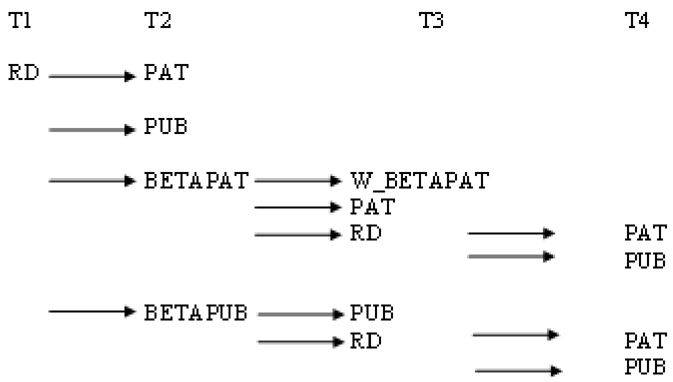
Table 4. Regression Results for Log (δ) for EU regions, 2001 (N=190)

Model	(1)	(2)	(3)
Estimation	OLS	OLS	ML – Spatial Error
Constant	-0.945***	-1.083***	-0.915***
	(0.060)	(0.068)	(0.190)
Log(GRD2000)	0.153***	0.130***	0.114***
	(0.010)	(0.012)	(0.009)
Log(PSTCK2001)		0.030***	0.008
		(0.008)	(0.009)
LAMBDA			0.936***
			(0.037)
R^2-adj	0.53	0.57	0.56
Multicollinearity Condition Number	7	9	
Jarque-Bera test on normality of errors	27	9	
Breusch-Pagan test for heteroskedasticity			38.830***
White test for heteroskedasticity	36.257***	45.754***	
LR-Err			86.716***
LM-Err			
Neighb	1.583	1.032	
INV1	74.335***	62.654***	
INV2	151.229***	133.821***	
LM-Lag			
Neighb	7.218***	5.809**	
INV1	63.334***	41.591***	0.006
INV2	151.780***	129.732***	0.032

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: LAMBDA is the spatial autoregressive coefficient; Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix. instruments in the IV-Spatial Lag estimation are W_Log(GRD2001), W_ Log(PSTCK2001), where W stands for the weights matrix INV2; *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

Finally Figure 3 summarises the dynamic effects of R&D on regional knowledge production suggested by our system of equations, and emphasises that the multiple feedback loops that operate in tandem.

Figure 3 – Spatio-temporal R&D policy effects



5. Concluding discussion and future plans

The current policy debate that advocates a spatial rationing of R&D resources may benefit by robust estimates on the likely effects of R&D expenditures in different types of EU regions. Since the current and forthcoming planning phases of the Framework Programme envisage a greater role for the regions, and an ever-greater proportion of the Structural Funds is being devoted to RTDI activities, the issue of the impact of spatial concentration at regional level is important for EU research and regional policy.

The results of the present study are still preliminary and should be treated with caution. If confirmed, our findings may have important policy implications. Preliminary findings indicate that there are distinct paths to 'critical mass' for Edison- and Pasteur-type research. The regional agglomeration of innovation capacity is important for Edison-type research whereas this is not the case for Pasteur-type research. Regional innovation capacity is built up slowly over time and comprises a 'knowledge' (knowledge stocks) element and an 'economy' element (knowledge-intensive employment). Using the modelling framework outlined above, we can estimate how the R&D productivity of European regions varies according to changes in local innovation capacity.

Our preliminary findings also indicate that networking (as promoted by the FPs) has been good at promoting Pasteur-type research. However, supporting European industry (the explicit aim of EU RTD policy) will probably have to go further. Should our results be confirmed, policy will need to prioritise improvements in local innovation capacity in order to improve R&D productivity. Our analysis here is in agreement with the description of regional knowledge production as a dynamic, positive-feedback process, where success breeds more success. In the near future we plan explore more closely the evolution of these effects over time.

Although one may be tempted to draw the conclusion that a spatial rationing of R&D would be an appropriate policy response, our results so far do not support this. Policy should be wary of redirecting R&D resources to the most productive regions for the very simple reason short-term gains in economies of scale may be long-term losses in economies of scope. Even if loses in variety were not concern, there are also indications that increases in R&D expenditures can have a detrimental effect on R&D productivity (Graves and Langowitz, 1996)[7]. Further research is needed to test the dynamic effects of changes in R&D inputs to R&D productivity.

It is likely though that different types of policies warranted for different types of European regions will be needed as well as a mixture of policy instruments ranging from education, corporate support and framework conditions requiring intensified policy coordination at various levels (European, national, regional).

In the near future, we plan to adjust our empirical analysis so as to more closely reflect the spatio-temporal R&D policy effects depicted in Figure 3. We also plan to use the modelling framework to perform an ex-ante evaluation of the likely effectiveness of FP support for research in different types of European regions. These can be gauged by simulations of in different "tiers" of EU regions according to beta categories. We may also simulate the likely effects of spatially varying distributions of the same aggregate FP6 expenditures.

[7] Recent research on a panel of developed economies, confirms this effect, and suggests that increases in R&D expenditures will have a negative effect on the R&D input/output ratio if they are sudden (or more specifically proportionately greater than the growth of existing knowledge stocks) (Pontikakis, 2008).

References

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

Acs, Z.J., Anselin, L. and Varga, A. (2002), "Patents and innovation counts as measures of regional production of new knowledge", *Research Policy*, Vol. 31, pp. 1069-1085

Andersen, E.S. (1992), “Approaching the National Systems of Innovation from the Production and Linkage Structure”, edited in Lundvall, B. Å. (ed), *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*, Pinter, London

Cooke, P. (2001), “Regional Innovation Systems, Clusters and the Knowledge Economy”, *Industrial and Corporate Change*, Vol. 10, No. 4, pp. 945- 974

Cooke, P. (2008), “Towards ‘Related Variety’ Ecologies”, Report for the Barcelona Workshop on *Research Specialisation in the EU*, organised by the EC DG JRC, Institute for Prospective Technological Studies (IPTS), 30 June (unpublished)

Edquist, C. (ed.) (1997), *Systems of Innovation*: *Technologies, Institutions and Organizations*, Pinter, London

Ellison, G. and Glaeser, E. (1997), “Geographic concentration in US manufacturing industries: a Dartboard approach”, *Journal of Political Economy*, Vol. 105, No 5, 889–927

Foray, D. and Van Ark, B. (2007), “Smart specialisation in a truly integrated research area is the key to attracting more R&D to Europe”, Policy Brief No 1, http://ec.europa.eu/invest-in-research/pdf/download_en/policy_brief1.pdf

Foray, D. (2008), “Understanding “Smart Specialisation””, Report for the Barcelona Workshop on *Research Specialisation in the EU*, organised by the EC DG JRC, Institute for Prospective Technological Studies (IPTS), 30 June (unpublished)

Georghiou, L. (2008), “Critical mass in the European Research Area”, Presentation at the conference on ‘Damaging Fragmentation or Healthy Diversity? – The Contribution of Economies of Scale to European Research’, Department for Universities, Innovation and Science (DIUS) London, 20 June

Graves, S.B. and Langowitz, N.S. (1996), “R&D Productivity: A Global Multi-Industry Comparison”, *Technological Forecasting and Social Change*, Volume 53, No. 2, pp. 125-137

Griliches, Z. (1990), “Patent statistics as economic indicators: A survey”, *Journal of Economic Literature*, Vol. 28, pp. 1661-1707

Hall, B.H. (1993), “The Stock Market's Valuation of R&D Investment During the 1980's”, *The American Economic Review*, Vol. 83, No. 2, pp. 259-264

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993), “Geographic localization of knowledge spillovers as evidenced by patent citations”, *Quarterly Journal of Economics*, 108, pp. 577–598

Jones, C. (1995), “R& D-Based Models of Economic Growth”, *Journal of Political Economy*, Vol. 103, No. 4, pp. 759-784

Maggioni, M.A, Nosvelli, M. and Uberti, T.E. (2007), “Space Vs. Networks in the Geography of Innovation: A European Analysis”, *Papers in Regional Science*, Vol. 86, No. 3, pp. 471-493

Malmberg, A. and Maskell, P. (1997), “Towards an explanation of regional specialization and industry agglomeration”, *European Planning Studies*, Volume 5, Number 1, pp. 25-41

Martin, B.R. (2008), “Critical Mass Effects in Research – The Concept and the Evidence”, Presentation at the conference on ‘Damaging Fragmentation or Healthy Diversity? – The Contribution of Economies of Scale to European Research’, Department for Universities, Innovation and Science (DIUS) London, 20 June

Mokyr, J. (2002), *The Gifts of Athena*, Princeton University Press, Princeton

Nelson, R.R. ed. (1993), National Innovation Systems: A Comparative Analysis, Oxford University Press, Oxford

Park, G. and Park, Y. (2006), "On the measurement of patent stock as knowledge indicators", *Technological Forecasting and Social Change*, Vol. 73, pp. 793-812

Ponds, R., van Oort, F.G., Frenken, K. (2008), “Internationalization and regional embedding of scientific research in the Netherlands”, edited in A. Varga (ed.), *Universities and Regional Development*: Edward Elgar, Cheltenham, UK and Northampton MA (in press)

Pontikakis, D., Chorafakis, G. and Kyriakou, D. (2008), “R&D Specialisation in the EU: from stylised observations to evidence-based policy”, IPTS Discussion Paper, Prepared for the Barcelona Workshop on *Research Specialisation in the EU*, organised by the EC DG JRC, Institute for Prospective Technological Studies (IPTS), 30 June (unpublished)

Pontikakis, D. (2008), “Economical innovation: measuring innovative productivity in a sample of developed economies”, Paper presented at the conference *Knowledge* *for Growth: European Strategies in the Global Economy*, French EU Presidency and Toulouse School of Economics, Toulouse, 7 July

Romer, P. M. (1990), "Endogenous Technological Change", *Journal of Political Economy*, Vol. 5, No. 98, pp. S71-S102

Stokes, D. E. (1997), *Pasteur's Quadrant: Basic Science and Technological Innovation*, Brookings Institution Press, Washington, D.C.

Varga, A. (2000), "Local Academic Knowledge Transfers and the Concentration of Economic Activity", *Journal of Regional Science*, Volume 40, No. 2, pp. 289-309

Varga, A. (2006), "The Spatial Dimension of Innovation and Growth: Empirical Research Methodology and Policy Analysis. *European Planning Studies* 9, 1171-1186

Varsakelis, N.C. (2006), "Education, political institutions and innovative activity: A cross-country empirical investigation", *Research Policy*, Vol. 35, pp. 1083-1090

Appendix 1 – Descriptive Statistics

(to be added)

Appendix 2 – List of regions

(to be added)