Journal of Economic Geography **14** (2014) pp. 229–263
Advance Access Published on 22 November 2012
doi:10.1093/jeg/lbs041

Metropolitan Edison and cosmopolitan Pasteur? Agglomeration and interregional research network effects on European R&D productivity

Attila Varga,†, Dimitrios Pontikakis** and George Chorafakis****

*Department of Economics and Regional Studies and MTA-PTE Innovation and Growth Research Group, Faculty of Business and Economics, University of Pécs, Pécs, Hungary
**Directorate for Science, Technology and Industry, Organisation for Economic Co-operation and Development (OECD), Paris, France
***Department of Geography, University of Cambridge, Cambridge, UK
†Corresponding author: Attila Varga, Department of Economics and Regional Studies and MTA-PTE Innovation and Growth Research Group, Faculty of Business and Economics, University of Pécs, Pécs, Hungary. *email* <vargaa@ktk.pte.hu>

Abstract

This article examines empirically the relative influence of static and dynamic agglomeration effects on the one hand and research networking [measured by Framework Programme (FP) participation] on the other on regional R&D productivity in the European Union. We found that agglomeration is an important predictor of R&D productivity in the case of market-oriented (Edison-type) research while interregional scientific networking is an important determinant of R&D productivity in the case of science-driven (Pasteur-type) research. Importantly, the two determinants are never *jointly* significant. This finding indicates that in a knowledge production context, and contrary to what may happen in other areas of economic activity, agglomeration and scientific networking are neither substitutes nor complements but operate at distinct parts of the knowledge production process. Our findings uncover the principal components of regional knowledge production processes across European regions in a dynamic setting. They therefore allow us to explore counterfactual scenarios and characterize the effects of policy interventions. A simulation of the likely impacts of FP6 funds on regional R&D productivity demonstrates that the dynamic effect is greater in regions with high agglomeration.

Keywords: Agglomeration, research networks, technology spillovers, patents, EU Framework Programs
JEL classifications: O31, R11
Date submitted: 24 June 2010 **Date accepted:** 2 October 2012

1. Introduction

A point of departure for this article is a seeming 'paradox' which has repeatedly drawn the attention of economists and economic geographers: on the one hand, regional economies tend to become increasingly interconnected and integrated in the global production of scientific and technological knowledge, as reflected in the increasing volumes of interregional collaboration in scientific publications, co-patenting, R&D joint ventures and other forms of inter-firm or academia-industry R&D collaboration, as well as in the intensified internationalization of R&D activities (Luukkonen et al., 1992; Caloghirou et al., 2004; EC, 2009). On the other hand, the production of scientific and technological knowledge is unevenly distributed in geographical space, as it tends to concentrate in a relatively small number of regional clusters which form the core of the global knowledge economy (e.g. Varga, 1999).

Generic studies of regional economies capturing this local–global duality are abundant in the economic geography literature. Regional economies of this type have been described, among others, as 'sticky places in a slippery space' (Markusen, 1996), 'Neo-Marshallian nodes of global networks' (Amin and Thrift, 1992) or clusters in which 'local buzz' coexists with 'global pipelines' (Bathelt et al., 2004). However, few studies quantify such phenomena in a knowledge economy setting (notable exceptions are Autant-Bernard et al., 2007; Maggioni et al., 2007) and to date no explicit comparison has been made between the relative importance of local and global effects from an econometric perspective. The need for a rigorous comparison in quantitative terms is brought to the fore by the intense policy debate on the optimal spatial allocation of EU research and innovation funding, dubbed 'smart specialization' (EC, 2010a, 2010b).

This article aims to contribute to this strand of literature by examining from an empirical point of view the effects of this local–global duality of regions on the knowledge economy, and more specifically, the co-existence of local, geographically mediated effects of agglomeration on the one hand, and of global, geographically non-embedded effects of networking on the other.

The main contributions of this article are in the following four aspects: first, it develops an integrated empirical model within which agglomeration and network effects on R&D productivity are jointly tested using European regional data. Second, the model distinguishes between science-driven and market-oriented scientific and technological research. This allows the impact of local/global effects to vary in each case and, therefore, to examine a richer set of possible contingencies. Third, the model considers both static and dynamic agglomeration effects, i.e. it also examines the cumulative impacts of R&D productivity on regional knowledge production. Fourth, the results of the empirical model are used to perform a policy impact analysis. Combined, these contributions yield insights that are of immediate relevance to contemporary (and recurrent) policy concerns, especially within the context of EU regional and research and innovation policy.

The second section of the article briefly presents the theoretical context of the main issues the article touches upon and the related literature; the third section introduces the empirical model; the fourth explains data and methodology; the fifth presents the empirical results; a policy simulation follows in the sixth section. The article concludes with a summary and some reflections on the policy implications of the analysis.

2. The theoretical context

2.1. Agglomeration effects

Agglomeration economies are external economies of scale, which emerge in geographical space. Marshall (1920) first distinguished between the traditional 'internal' economies of scale, coming from the expansion of the scale of operation of a firm and 'external' economies induced by spatial proximity, which arise from the expansion of whole industries. Intra-industry, spatially concentrated, 'Marshallian' externalities are known as 'localization economies'; inter-industry externalities, also mediated by geographical space, are known as 'urbanization economies'. Glaeser et al. (1992)

distinguish between a 'Marshall-Arrow-Romer' type of externalities caused by intra-industry, usually vertical knowledge spillovers within the same value chain and a 'Jacobs' (1969) type of externalities, caused by inter-industry, horizontal knowledge spillovers between parallel value chains; the former is a dynamic form of localization while the latter of urbanization economies.

Agglomeration externalities are thought to be induced by labour pooling or more generally the localized accumulation of human capital, the emergence of 'untraded interdependencies', informational externalities (Dosi, 1988; Storper, 1997) and trust or more generally the accumulation of social capital and the density of markets for intermediate products and outputs. Agglomeration economies are widely recognized as being capable of increasing firms' productivity via several different routes; empirical studies have demonstrated both the direct causal effects of agglomeration on firm productivity, as well as its indirect effects through wages, firm birth or employment (Rosenthal and Strange, 2004).

Innovation and, consequently, R&D investment are commonly considered as key factors for increasing the productivity of firms, as well as of regional and national economic systems. The effect of agglomeration economies on the innovative capacities of firms or of entire economic systems and, in particular, on the regional knowledge production process, is a factor which has been taken into account, albeit tangentially, in several empirical studies (examples include Jaffe, 1989; Audretsch and Feldman, 1996; Anselin et al., 1997; Crescenzi et al., 2007). However, an explicit analysis of the role agglomeration plays in the efficient deployment of R&D in regional economies still remains an underexplored topic. Among the exceptions is Varga (2000, 2001), who tests econometrically in a knowledge production function (KPF) setting the role of agglomeration in the R&D productivity of universities using data on US metropolitan statistical areas. The study finds the existence of a 'critical mass' of advanced technology firms, private research labs and business services directly associated with a sizable labour pool in the urban high-technology sector as being a prerequisite for a significant impact of university R&D on regional innovation. Further studies in this strand include Koo (2005), who developed an endogenous approach, Acs and Varga (2005) on the roles of agglomeration and entrepreneurship in Europe and Goldstein and Drucker (2006) on the impact of city size on regional economic roles of US universities. It is also worth mentioning here Feldman (1994), who brought attention from a more qualitative and case-specific point of view to the then suboptimal regional role of the Johns Hopkins University in transferring knowledge to the local economy. The study points to the relatively underdeveloped technology sector in the region as perhaps the main reason of this anomaly. This case suggests that even a university with outstanding research activity is not capable of transferring substantial knowledge to the local economy without a concentration of innovative firms and private research labs ready to absorb that knowledge or business services participating in the various stages of the innovation process.

2.2. Network effects

The properties and effects of social networks have been studied extensively from various perspectives. This emerging methodological field in the social sciences was explored by sociologists and anthropologists (e.g. Granovetter, 1973; White, 1992), as well as mathematicians and physicists (e.g. Barabási and Albert, 1999; Newman, 2000), long before the important effects of networking on fundamental economic processes drew the attention of economists and economic geographers. More recently the realization of the essential role of networks in the learning process of economic agents and in particular of the firms, in the formation of inter-firm strategic alliances and the accumulation of social capital, and finally—and probably most importantly—in the diffusion of knowledge spillovers, the generation of scientific and technological knowledge and, consequently, the innovation process, has led to a proliferation of papers in economics and economic geography on theoretical and empirical aspects of knowledge networks.

A strand of this literature approaches specific aspects of knowledge, innovation and R&D networks from a theoretical perspective, often in a game-theoretical setting. Examples include various stylized models of inter-firm network formation through strategic R&D collaboration and search for knowledge spillovers (Goyal and Moraga-Gonzalez, 2001; Cowan, 2004; Andergassen et al., 2005; Cowan and Jonard, 2006). Other papers examine theoretically inter-firm networks and their innovative performance from the perspective of strategic management (Hite and Hesterly, 2001; Stuart and Sorenson, 2007). A different strand of network literature focuses from an empirical perspective on the structure and properties of specific types of knowledge networks, notably research collaboration networks such as co-patenting (Balconi et al., 2004; Carayol and Roux, 2007); co-authorship (Newman, 2001; Wagner and Leydesdorff, 2005; Fafchamps et al., 2006) and EU Framework Programme (FP) collaboration networks (Barber et al., 2006; Billand et al., 2008). Some papers specifically focus on the role of networks in the transmission of scientific and technological knowledge from academia to industry; Varga and Parag (2009), for example, examine the impact of the co-publication network structure on university patenting.

Finally, an increasing number of studies approach the issue from a spatial perspective, where 'spatial' should be interpreted both in the context of physical and 'relational' space, focusing on the distinct effects of geographical and relational proximity. Johansson and Quigley (2004) compare from a theoretical perspective the parallel developments in the economics of agglomeration and of networks, arguing for the substitutability of agglomerations by networks. Gastner and Newman (2006) model geographically embedded networks and examine their costs and benefits. Breschi and Lissoni (2005) test the existence and magnitude of localized knowledge spillovers by using patent data to control for the mobility of inventors across companies and space, to conclude that access to local pools of knowledge is not ensured by mere geographical proximity but requires active participation in knowledge exchange networks. Ponds et al. (2007, 2009) analyze the role of geographical proximity for collaborative scientific research between universities, firms and public research institutes using co-publication data and demonstrate that collaboration between different kinds of organizations is more geographically localized than collaboration between organizations that are similar due to institutional proximity. Maggioni et al. (2007) examine the relative significance of geographical and relational spillovers among European regions for their innovative capacities by econometrically comparing participation in two research networks, namely those of FP5 and of EPO co-patent applications; the main idea of the article is that knowledge is created when crucial actors co-locate in geographical space, thus giving birth to regional clusters, industrial districts, excellence centres, etc. and is subsequently diffused either due to spatial contiguity or through a-spatial networks.

Autant-Bernard et al. (2007) examine to what extent network and geographical effects are determinants of collaboration along with other microeconomic factors using FP6 participation data, to conclude that the probability of collaboration is influenced by the individual's position in the network and that social (i.e. relational) distance matters probably more than geographical distance. The present article belongs to this last strand of literature.

The causal links between the degree of connectedness and innovativeness, productivity and competitiveness of firms and regions are relatively well documented. These causal relations make possible, at least in theory, that even regional economies which exhibit weak agglomeration effects but are well embedded in global knowledge production networks be highly productive; this means that increasing interregional connectedness may be an alternative explanation of regional R&D productivity to agglomeration economies.[1]

The attempts we have documented so far to gauge the effects of agglomeration and networking either separately or in combination, map such effects to a rather narrow set of outcomes (collaboration, innovation, cluster location) that are difficult to link together for comprehensive policy evaluation. Moreover, to date, no study weighs the impact of interregional connectedness *vis-à-vis* agglomeration on *R&D productivity* in particular. In light of the significance of R&D productivity for the long-term growth of knowledge-based economies, this is an important gap in the current literature which this article hopes to partially fill.

2.3. Types of knowledge and types of research

Much of the knowledge required in the production of new technologies is tacit, that is, knowledge obtained by experience, embodied in individuals and diffusing primarily by way of interpersonal contact. In a knowledge production setting, proximity to places with a high concentration of people possessing this type of knowledge becomes crucial. By contrast, the diffusion of codified knowledge is generally not conditional on proximity. Modern ICTs facilitate its diffusion and arguably the intensity of its use in knowledge production, to a greater extent than ever before. Indeed, the importance of locally contained knowledge in the formation of geographical clusters is well documented (Audretsch and Feldman, 1996). As demonstrated by patent citations, for certain types of technological knowledge, diffusion is highly concentrated geographically (Jaffe et al., 1993).

Importantly, different types of research impose different requirements on scale and place a different emphasis on tacit knowledge and by extension, on proximity (Malmberg and Maskell, 1997). Taking into account the sharp institutional differences in the worlds of scientific and technological research and using the terminology introduced by Stokes (1997), we consider two distinct types of research:

(a) *Edison*-type: research, whose products have clear economic applications, pursuing market-oriented innovation;

1 Furthermore, even agglomeration phenomena can be interpreted as a localized type of network effects. In this context, agglomerated knowledge production systems can also have a network representation and agglomeration effects can be interpreted from a network-analytical perspective. It can be further argued that the type of knowledge that is critical for a particular economic system determines the structure of its network representation.

(b) *Pasteur*- (and implicitly Bohr[2]-) type: science-oriented research, mediated by the distinct rules and incentives of the scientific establishment—sometimes dubbed 'pre-competitive research' among EU policy analysts (and referred as such in relevant EU treaties).

Given the different spatial diffusion dynamics of tacit and codified knowledge and the relative importance of tacit knowledge for Edison-type research, a preliminary hypothesis can be sketched: the prevalence of agglomeration over network effects (and vice versa) may correspond to qualitative differences in the type of research involved and its respective knowledge input requirements. To investigate such differences, our empirical analysis examines separately agglomeration and network effects for Edison- and Pasteur-type research.

2.4. Policy relevance

Besides its self-standing analytical value, the central question posed by this article is of high relevance to ongoing discussions on the future directions of EU research and innovation policy. A recurrent issue in EU policy discourse is the optimal geographical and sectoral allocation of resources for research (see contributions to Pontikakis et al., 2009, especially by Foray and Cooke, and EC, 2010b; for earlier accounts from an industrial/technology policy perspective see Geroski, 1989a, 1989b; Matthews and McGowan, 1992). This stems from a concern that EU research funds are spread too thinly across Europe without achieving economies of scale that would strengthen the overall competitiveness of the EU *vis-à-vis* its main technological and economic rivals, and without attaining the impact on growth and employment that is expected from them. A policy-induced geographical and sectoral concentration of R&D resources on the basis of existing patterns of technological specialization, coined 'smart specialization', is put forward as one possible solution to the perceived problem (Foray and van Ark, 2007; Foray, 2009; McCann and Ortega-Argilés, 2011).

An alternative policy prescription to the induced concentration of R&D resources is to promote cross-regional research networks connecting complementary research capabilities not available within own regions (Chorafakis and Pontikakis, 2011). A policy of sustaining or even increasing the degree of connectedness in EU research or 'networked specialization' is therefore suggested as a possible alternative policy option (Georghiou et al., 2008).

So far, this debate rests on scattered sources of empirical evidence and lacks a comprehensive approach. By developing and testing an empirical model that considers the effects of both agglomeration and networking on R&D productivity, this article provides a framework within which alternative policy suggestions can be weighed against each other.

2 Following Mokyr (2002), we narrow down Stokes' (1997) three types to just two: as our concern is with economically useful knowledge, the distinction of importance is between R&D motivated primarily by a quest for fundamental understanding versus knowledge primarily motivated by profit [c.f. 'propositional' versus 'prescriptive' knowledge, in Mokyr (2002)].

3. The empirical modelling framework[3]

Our starting point is the KPF initially specified by Romer (1990) and parameterized by Jones (1995). In the interpretation of the parameters, we follow Varga (2006).

$$\mathrm{d}A_i/\mathrm{d}t = \delta H_{Ai}^{\lambda} A_i^{\varphi}, \tag{3.1}$$

where $\mathrm{d}A/\mathrm{d}t$ is the temporal change in technological knowledge, H_A refers to research inputs (e.g. number of researchers or research expenditures), A is the total stock of already existing scientific and technological knowledge (knowledge codified in publications, patents, etc.) and i is the index of the spatial unit. In Equation (3.1), technological change is associated with contemporary R&D efforts and previously accumulated knowledge. The same number of researchers can have a varying impact on technological change depending on the stock of already existing knowledge. Two parameters in Equation (3.1) are particularly important for this article: the size of φ reflects the impact of codified knowledge transfer. Since codification makes knowledge diffusion possible over large distances, this parameter reflects knowledge flows with unlimited spatial accessibility. Regarding the parameter λ, the larger its size, the stronger the impact the same number of researchers plays in technological change. Its value reflects (codified and tacit) knowledge transfers within the research sector and between the research sector and the rest of the innovation system. The innovation literature highlights the importance of interactions among the various actors in the innovation process (e.g. Nelson, 1993; Edquist, 1997). Thus, knowledge transfer depends on the intensity of interactions among researchers (H_A); the size and quality of public research; the extent to which the private research sector interacts with it (especially with universities) via formal and informal linkages; the development level of supporting/connected industries and business services and the integration of innovating firms into the system via linkages to them (Andersen, 1992; Cooke, 2001). Therefore, the characteristics of the broader innovation system play a key role in the productivity of research, as reflected in the size of λ.

Some of the interactions of researchers are localized, especially those that require tacit knowledge transfers or frequent connections in collaboration, whereas others can be maintained over larger distances via for example formal research network linkages. The size of λ is positively related to the concentration of innovation system actors in the proximity of research labs on the one hand and to the intensity of interactions through interregional research networks on the other. Thus, we assume that both agglomeration and interregional research networking strengthen regional research productivity.

The theoretical and empirical literature on economic geography has highlighted the cumulative, self-reinforcing nature of agglomeration (e.g. Fujita et al., 1999; Fujita and Thisse, 2002). In our modelling framework, we assume that agglomeration of innovation system actors and resources also occurs in a cumulative, dynamic fashion. Research productivity (resulting either from agglomeration or from interregional research networking or from both) can be a revealing summary measure of a regional innovation system's qualities. Therefore regions with high research productivity act as centres of gravity for further research resources; private R&D activities are attracted by expectations of high returns, as are greater portions of competitive public research

3 This section draws extensively on Varga (2006).

funding. Increased research activities may then cause increased agglomeration by drawing further actors to the region (such as innovative firms or specialized business services) with the expectation of emerging opportunities in innovation. Thus, we hypothesize that a gradual self-reinforcing process shapes the geographical structure of innovation.

The extent to which the processes described above work is not yet known. To the best of our knowledge, this article represents the first attempt to empirically investigate the role of static and dynamic agglomeration and interregional networking on research productivity. We test our hypotheses with a four-equation empirical model. This model is the extension of the static analysis developed and applied in Varga (2000, 2001).

In order to test empirically the hypothesized relationships, we use the following econometric specifications: using subscripts i and N to denote individual regions and nations (in our case EU member states), respectively, the empirical counterpart of the Romer KPF[4] is specified as:

$$\mathrm{Log}(K_i) = \alpha_0 + \alpha_1 \mathrm{Log}(\mathrm{RD}_i) + \alpha_2 \mathrm{Log}(\mathrm{KSTCK}_N) + \varepsilon_i, \tag{3.2}$$

where K stands for new scientific-technological knowledge, RD is expenditure in research and development and KSTCK represents existing technological knowledge at the national level. We use the national patent stock as a proxy for codified technological knowledge reachable with unlimited spatial accessibility within the country.

Equation (3.3) relates research productivity measured by $\alpha_{1,\mathrm{i}}$, the parameter of the research variable in Equation (3.2), to agglomeration and interregional networking.

$$\alpha_{1,i} = \beta_0 + \beta_1 \mathrm{Log}(\mathrm{AGGL}_{i,t-k}) + \beta_2 \mathrm{Log}(\mathrm{NET}_{i,t-\mathrm{k}}) \tag{3.3}$$

where AGGL_i measures the agglomeration of innovation system actors in the region and NET is for interregional research networks.

Substituting Equation (3.3) into Equation (3.2) results in the following equation to be estimated:

$$\begin{aligned} \mathrm{Log}\big(K_{i,t}\big) = \alpha_0 + \beta_0 \mathrm{Log}(\mathrm{RD}_{i,t-\mathrm{k}}) + \beta_1 \mathrm{Log}(\mathrm{AGGL}_i) \times \mathrm{Log}(\mathrm{RD}_{i,t-\mathrm{k}}) \\ + \beta_2 (\mathrm{NET}_{i,t-\mathrm{k}}) \times \mathrm{Log}(\mathrm{RD}_{i,t-\mathrm{k}}) + \alpha_2 \mathrm{Log}(\mathrm{KSTCK}_{N,t-\mathrm{k}}) + \varepsilon_i, \end{aligned} \tag{3.4}$$

Following on, in order to test the cumulative nature of agglomeration, the determinants of location of R&D expenditures ($\mathrm{RD_i}$) and of innovation actors can be empirically modelled by:

$$\mathrm{d}(\mathrm{RD}_{i,t}) = \lambda_0 + \gamma_1 \alpha_{1,i,t-k} + \lambda_1 \mathrm{Z}_{1,i,t-k} + u_i \tag{3.5}$$

$$\mathrm{d}(\mathrm{AGGL}_{i,t}) = \xi_0 + \xi_1 \mathrm{RD}_{i,t-k} + \xi_2 \mathrm{Z}_{2,i,t-k} + \mu_i \tag{3.6}$$

where variable Z_1 and Z_2 stands for additional control variables.

This framework allows to test various alternative hypotheses.

4 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 lessening the influence of outliers and allowing for direct comparisons of coefficients for variables expressed in different units of measurement.

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, following the terminology concerning the different types of scientific and technological research presented in the introduction, we observe that 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 KPFs to draw our comparisons.

4. Data and estimation issues

Our empirical analysis is based on a sample of 189 European regions (a mixture of NUTS2 and NUTS1 regions) where information was complete enough for the purposes of our study (see Table A2 for a list of regions). We use a mixture of panel [for the KPFs, i.e. Equation (3.4)] and cross-sectional analysis [for the temporal change of R&D and employment equations, i.e. Equations (3.5) and (3.6)] depending on the nature of the underlying question and data availability.

The time period under examination is determined by the duration of the EU 5th Framework Programme (FP5) spanning the years 1998–2002, as our measure of interregional networking draws on administrative data from this particular policy instrument. To reflect the interval between the performance of R&D and its translation into measurable outputs, the independent variables are lagged. There is no agreement in literature as to the ideal duration of a lag and attempts to estimate it empirically have been inconclusive (Hall et al., 1986). In practice, aggregate studies of KPFs with patents commonly employ 2- or 3-year lags (Furman et al., 2002; Furman and Hayes, 2004). Our own experimentation with lags of varying duration showed that they produce very similar results.[5] Temporally lagged-dependent variables have the added advantage of lessening the potential for endogeneity problems. We therefore opted for the theoretically plausible 2-year lag. The combination of the boundaries set by the duration of FP5 and the 2-year lag mean that our panel runs for the 3-year period 2000–2002 (1998–2000 for the independent variables). A summative description of the variables used in the study and the data sources can be found in Table 1 (descriptive statistics in Table A1).

Further to this concise description, a few additional words of clarification regarding the choice, construction and limitations of the variables are in order. We use patent applications to the EPO ($PAT_{i,t}$) and scientific publications in ISI journals ($PUB_{i,t}$) as proxies for Edison- and Pasteur-type knowledge flows, respectively. Although patent counts are far from a perfect proxy of innovation (e.g. among other things, not all innovations are patentable or patented, for a comprehensive assessment see Griliches, 1990), the patent examination process and the cost it implies for applicants, present a more or less objective yardstick of substantial novelty. Moreover, patents are the only measure that is available for a large number of European regions and over a number of years. The 'law of large numbers' (Griliches, 1990) provides a justification for their use,

5 This result repeats what is experienced with US data in a similar KPF context (Varga et al., 2005).

Table 1. Variables used in the study

Variable name	Description	Source
$PAT_{i,t}$	Number of patent applications to the European Patents Office (EPO) by region of inventor, sorted by date of application (priority year). Fractional counts.	Eurostat NewCronos database
$PUB_{i,t}$	Number of publications in scientific journals in the Thomson ISI database (search criteria: article, letter, review)	RKF database (data processed by CWTS, Leiden University)
$GRD_{i,t}$	Gross regional expenditures on R&D, in millions of Purchasing Power Standard (PPS) Euros. 1995 prices.	Eurostat NewCronos database
$KSTCK_{N,t}$	National patent stocks for the five previous years, depreciated by 13% (PIM).	Authors' elaboration of Eurostat NewCronos
$EMPKI_{i,t}$	Employment in technology and knowledge-intensive sectors. Measured in thousands of people.	Eurostat NewCronos database
$\delta_{i,t}$	Index of agglomeration. Size-adjusted location quotient of employment in technology and knowledge-intensive sectors.	Authors' elaboration of Eurostat NewCronos
$NETGRD_{i,t-k}$	Total of the (log of) R&D expenditures in network partner regions for each region as a proxy for interregional network effects.	Authors' elaboration of FP5 administrative database, DG RTD, Dir A
$PUBCORE_i$ $RDCORE_i$	Dummies taking a value of 1 for regions with a number of publications (PUBCORE)/gross R&D expenditures (RDCORE) >1 SD from the sample mean, zero otherwise.	Eurostat NewCronos database
$PATHCORE_i$ $RDHCORE_i$	Dummies taking a value of 1 for regions with a number of patents (PATHCORE)/R&D expenditures (RDHCORE) >2 SDs from the sample mean, zero otherwise.	Eurostat NewCronos database
$ALPHAPAT1998_i$	R&D productivity estimates for Edison-type knowledge (patents) across European regions controlling for other factors. 1998 values. Corresponds to coefficient α_{1i} in Equation (3.2).	Authors' estimates
$ALPHAPUB1998_i$	R&D productivity estimates for Pasteur-type knowledge (publications) across European regions controlling for other factors. 1998 values. Corresponds to coefficient α_{1i} in Equation (3.2).	Authors' estimates
DGRD01-98	Temporal change in R&D expenditures over the period 1998–2001. ($=GRD_{i,2001} - GRD_{i,1998}$).	Eurostat NewCronos database
DEMPKI01-98	Temporal change in employment in technology and knowledge-intensive sectors over the period 1998–2001. ($=EMPHT_{i,2001} - EMPHT_{i,1998}$).	Eurostat NewCronos database
WAGES	Wage income/employment, 1998	Eurostat
POPDENS	Population/area, 1998	Eurostat
MARKETPOT	Sum of inverse distance weighted GDP in all NUTS two regions	Eurostat
$NORMCIT_{N,t}$	Normalized citations. Calculated as the ratio of the national citations per paper to the world citations per paper.	SCImago (2012)

especially, we may add, for large spatial units.[6] Comfortingly, previous research has shown that at the level of regions, patent counts correlate well with innovation counts (Acs et al., 2002) and both measures provide very similar results in the KPF context. Likewise, the number of journal publications is a commonly used indicator of scientific output (van Raan, 2004; Azagra-Caro et al., 2007; Crespi and Geuna, 2008). Publications are, arguably, a somewhat stronger proxy (as compared to patents) for the 'true' amount of (in their case, Pasteur-type) knowledge flows, given the *de facto* requirement to publish the results of scientific R&D. Such bibliometric indicators though are not without problems[7] themselves, including the possibility of bias in journal coverage and the distorting effects of evaluation mechanisms.[8] In the regression context, we control for potentially important differences across countries in the quality of scientific output using an indicator of national scientific impact drawing on publication citation data. The rationale is that higher quality publications will, on an average, have a greater impact in terms of citations in subsequent publications. Our specific variant, the number of citations per paper, normalized for the world citation rate ($\mathrm{NORMCIT}_{N,t}$) follows OECD (2012, 47) practice.

Following Romer (1990), the importance of knowledge stocks (or a 'standing on the shoulders of giants' effect) for knowledge production has been verified empirically (Furman et al., 2002; Zucker et al., 2007). Three different types of national patent stocks were constructed and tested empirically: patent stocks with no depreciation (Porter and Stern, 2000; Furman et al., 2002) and, using the perpetual inventory method (PIM), patent stocks with a 13 (Park and Park, 2006) and 15% 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:

$$\mathrm{PSTD}_{N,t} = \mathrm{PSTD}_{N,t-1} \times (1-d) + \mathrm{PAT}_{N,t}$$

where d is the depreciation rate (13 or 15%). Initial stocks take into account compound annual growth in the five preceding years.[9] After testing all three variants and observing

6 Invoking this assumption of course implies sidelining the important issue of patent quality or the common observation that the economic value of patents is highly skewed: Insofar as we are concerned with the knowledge-generating sector and are not drawing inferences about the economy at large, this issue lies outside the scope of the present article.

7 It is worth noting that CTWS University of Leiden (the ECs data provider for this metric in the RKF database) performs a 'field normalization' to account for the uneven propensity to publish across disciplines along with several other quality adjustments documented in Moed et al. (1995) and van Raan (2004).

8 As pointed out by one anonymous reviewer, the presence of formal evaluation mechanisms in countries such as UK and the Netherlands raises plausible concerns of inflationary tendencies and a systematic bias in part of our sample. However, the fact that these countries also typically score very highly in terms of citations, both in terms of general citation counts (May, 1997, 793) and in terms of the top 10% of cited publications (EC, 2011, 139) suggests that the real value of publications originating there is not negatively affected.

9 Initial stock equals flows for first year divided by the sum of compound growth for the preceding 5-year period and the depreciation rate. 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 2 years), while the 1998 value was estimated as the average of 1997 and 1999.

that results do not differ, our final estimates use the PIM stocks with a 13% depreciation rate ($KSTCK_{N,t}$).

The region's level of agglomeration δ is proxied by a novel index of agglomeration of knowledge intensive employment. As most measures of absolute concentration of economic activity introduce multicollinearity, they are likely to be problematic in a regression context with interaction terms. Our index is a size-adjusted [in the spirit of the index developed by Ellison and Glaeser (1997)] variation of the popular location quotient (LQ) measure and is calculated as follows:

$$\delta_i = [(\mathrm{EMPKI}_i/\mathrm{EMPKI}_{\mathrm{EU}})/(\mathrm{EMP}_i/\mathrm{EMP}_{\mathrm{EU}})]/[1 - \sum_j (\mathrm{EMPKI}_{i,j}/\mathrm{EMPKI}_{j,\mathrm{EU}})] \times [1 - (\mathrm{EMP}_i/\mathrm{EMP}_{\mathrm{EU}})],$$

where EMPKI_j and EMPKI are employment in knowledge intensive economic sector j and the total of knowledge intensive sectors,[10] EMP is total employment and the subscripts i and *EU* stand 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 penalize 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 for the agglomeration of innovation system actors.

With respect to our measure of interregional scientific networking, we derive it from the European Commission CORDA database of participations in FP5. There are good reasons to expect that participations to the FP can be an appropriate proxy of the relational structure of interregional knowledge diffusion across Europe. The FPs were designed to support 'pre-competitive', collaborative research with no national bias as to the types of technologies promoted and the distribution of funds. The pre-competitive character of supported research ensured that Community funding did not clash with the competition principles of the Common Market and did not function as a form of industrial subsidy; the collaborative character of research and the cost-sharing provisions were seen to guarantee the diffusion of technologies and the involvement of various types of actors from the whole technological knowledge creation spectrum, such as large and small firms, universities and public research institutes. One potential drawback of the FP as a data source is the fact that it is artificial; i.e. collaborating teams will not always coincide with naturally occurring networks of researchers. However, at as an aggregate level as that of a region and given the FPs overall gravity in European research,[11] differences between the two are arguably negligible.

10 The classification of knowledge intensive economic sectors (devised by Eurostat) includes: high and medium high technology manufacturing, 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).

11 According to EC (2009, 103), European funding accounted for 12–15% of public R&D expenditures in Europe over the period 1995–2006, of which about half is channelled through the FP. The total amount of funding allocated to FP collaborative projects is twice as large if one also takes into account co-funding from national sources (Barré et al., forthcoming).

Using the FP5 database, we have constructed an n by n matrix (where n = number of NUTS 1 and 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 total of the (log) R&D expenditures in network partner regions for each region as a proxy for interregional network effects ($\mathrm{NETGRD}_{i,t-k}$).

Tests for pooling, multicollinearity, heteroskedasticity, spatial dependence and endogeneity are run and, where appropriate, adjustments are made in the estimations.

5. Empirical results

Following the equations specified in Section 2, we first estimate the KPF using patents as a proxy of Edison-type knowledge across European regions over the 3-year period 2000–2002 (Table 2). Regressions were estimated in 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.[12] The first baseline model 1 confirms that, on an average, lagged gross regional R&D expenditures (GRD) 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 an average about proportionate to R&D inputs.

Model 2 includes the product of lagged R&D expenditures and δ. Model 2 suggests that agglomeration 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 innovation (Andersen, 1992; Nelson, 1993; Edquist, 1997; Cooke, 2001). The importance of co-location is also suggestive of the significance of tacit knowledge (Malmberg and Maskell, 1997).

Model 3 tests the significance of research network effects, by including the product of gross R&D expenditure of region i times the (logarithm of the) value of the sum of R&D expenditures of those regions with which region i had at least one joint research project in FP5 [$\mathrm{Log(GRD)} \times \mathrm{NETGRD}_{t-2}$]. The product term is statistically insignificant. This result suggests that R&D expenditures of collaborating regions do not affect R&D productivity in the region.[13]

Model 4 introduces national patent stocks (KSTCK), indicating that historically accumulated technological knowledge has a positive, statistically significant and quantitatively distinct effect on regional patenting. Interestingly, the coefficient of $\mathrm{Log(GRD)} \times \mathrm{Log}(\delta)$ drops from ~0.32 in models 2 and 4 to ~0.24, suggesting that codified knowledge spillovers capture at least some of the effects attributed to agglomeration in the previous models. In models (1–5), the LM-tests confirm the presence of a strong spatial dependence even after controlling for model variables.

12 The multicollinearity condition number is the square root of the ratio of the largest to the smallest eigenvalue of the matrix X'X after standardization. As a rule of thumb values of the condition number exceeding 30 signals serious multicollinearity (Belsley et al., 1980).

13 Of course, this does not conclusively disprove the existence of interregional network effects (possibly by other means) not captured by our coarse proxy.

Table 2. Regression results for log(patents) for 189 EU regions, 2000–2002 ($n = 567$)

Model Estimation	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 2SLS-spatial lag (INV2)
Constant	−1.6421*** (0.1776)	−0.3107 (0.2316)	−0.5391* (0.2806)	−1.7864*** (0.2381)	−1.7227*** (0.2372)	−2.3006*** (0.2743)
W_Log(PAT)						0.2455*** (0.0631)
Log(GRD$_{t-2}$)	1.0822*** (0.0308)	0.8453*** (0.0407)	0.9585*** (0.0886)	0.7142*** (0.0377)	0.6879*** (0.0384)	0.7088*** (0.0377)
Log(GRD$_{t-2}$) × Log(δ_{t-2})		0.3242*** (0.0389)	0.3222*** (0.0389)	0.2443*** (0.0351)	0.2136*** (0.0363)	0.1439*** (0.0396)
Log(GRD$_{t-2}$)*			−8.675E−05 (6.03E−05)			
NETGRD$_{t-2}$				0.2502*** (0.0203)	0.2536*** (0.0202)	0.1804*** (0.0272)
Log(KSTCK$_{t-2}$) PAHTCORE					0.4814*** (0.1568)	0.4614*** (0.1526)
R^2-adj	0.69	0.72	0.72	0.78	0.78	
Log likelihood	−885.30	−852.36	−851.32	−784.69	−779.98	
Sq. corr.						0.80
Multicollinearity condition number	7	10	24	13	13	
F on pooling (time)	0.9071	0.6777	0.5644	0.8143	0.6425	
F on slope homogeneity	0.4815	0.7613	0.5836	0.6485	0.4645	
White test for heteroscedasticity	0.7529	1.0462	12.8409	3.6634	12.1852	
LM-Err						
Neigh	111.78***	69.36***	66.85***	26.95***	23.46***	
INV1	252.17***	129.64***	117.26***	29.87***	26.13***	
INV2	215.12***	121.59***	114.45***	32.40***	29.24***	
LM-lag						
Neigh	142.53***	100.88***	99.03***	24.99***	25.89***	
INV1	247.03***	159.07***	153.47***	28.16***	27.96***	
INV2	237.99***	148.93***	145.48***	31.42***	30.95***	

Estimated SEs 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(PAT) is the spatially lagged dependent variable, where W stands for the weights matrix INV2. ***indicates significance at $p<0.01$; **indicates significance at $p<0.05$; *indicates $p<0.1$. In model 6, the Durbin–Wu–Hausman test for Log(GRD$_{t-2}$) and Log(GRD$_{t-2}$) × Log(δ_{t-2}) does not reject exogeneity. The instruments were selected following the three-group method. For the spatial lag term, the instruments are the spatially lagged explanatory variables.

Though the explanatory variables lag 2 years behind the dependent variable and as such no endogenous relationship is expected in the equation, stability in the spatial structure of R&D in a medium term might be the source of correlation between the explanatory variables and the error term. However, the Durbin–Wu–Hausman test does not reject exogeneity for the regional left hand side variables.[14]

Spatial lag dependence captured by the square inverse distance matrix is the most significant, thus it is used in the final estimated model.[15] The significant coefficient of the spatially lagged-dependent variable [W_log(PAT)] indicates that regional patenting activity is positively affected by interregional knowledge flows. However, these flows are not mediated by formal research collaborations as shown by the insignificant network effect in model 3 but by interactions fading away with distance. Given that error terms are not distributed normally, the appropriate regression is the spatial lag model estimated with the instrumental variables methodology (2SLS). In model 6, controlling for spatial dependence, the substantive results remain unaffected, although the value of the coefficient for the agglomeration interaction term is smaller. The dummy variable PATHCORE (with 1 for regions with >2 SDs above the EU average patent applications and 0 otherwise) enters the equation with significant coefficients in models 5 and 6 suggesting remarkable differences between high and low patenting regions in Europe. It is worth noting that all models explain 70% or more of the variation in regional patenting.

Table 3 estimates the KPF with scientific publications as the dependent variable. In all models, regression diagnostics indicate no problems with multicollinearity and, as with patents, the KPFs explain >70% 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 [included either with or without the cross-product variable $\text{Log}(\text{GRD}_{t-2}) \times \text{NETGRD}_{t-2}$ as it is in Model 3], while research network effects (Models 2–6) exert a statistically significant and quantitatively distinct influence on scientific R&D productivity.

Therefore, in the case of Pasteur-type research, interregional scientific networking is more important than local agglomeration. In other words, regions can perform research efficiently even in the absence of local agglomeration. 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. No significant spatial dependence is found but heteroscedasticity remains persistently present throughout the models. Given

14 The three-group method suggested by Kennedy (1998) was followed in instrument selection. For each variable the instrument takes the value −1, 0 or 1 according to whether the value of the instrumented variable is in the lower, middle or upper third of its ranking.

15 The general expression for the spatial lag model is:

$$y = \rho Wy + x\beta + \varepsilon,$$

where y is an N by 1 vector of dependent observations, Wy is an N by 1 vector of lagged-dependent observations, ρ is a spatial autoregressive parameter, x is an N by K matrix of exogenous explanatory variables, β is a K by 1 vector of respective coefficients and ε is an N by 1 vector of independent disturbance terms. Because the spatially lagged-dependent term is correlated with the errors, the OLS estimator is biased and inconsistent. Instead of OLS, other estimation methods such as Maximum Likelihood, Instrumental Variables or General Methods of Moments must be applied to the spatial lag model (Anselin, 1988).

Table 3. Regression results for log(publications) for 189 EU regions, 2000–2002 ($n = 567$)

Model Estimation	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 2SLS Hetsced robust	5a OLS (robust-ness check)
Constant	1.4026*** (0.1298)	2.3886*** (0.1645)	2.196*** (0.202)	2.3395*** (0.1711)	2.4568*** (0.1697)	2.6137*** (0.3199)	2.4122*** (0.1824)
$Log(GRD_{t-2})$	0.942*** (0.0225)	0.445*** (0.0597)	0.480*** (0.633)	0.4158*** (0.066)	0.4523*** (0.0602)	0.4317*** (0.1262)	0.4602*** (0.0614)
$Log(GRD_{t-2}) \times Log(\delta_{t-2})$			−0.0462 (0.0282)				
$Log(GRD_{t-2})$* $NETGRD_{t-2}$		0.0004*** (4.4E − 05)	0.0004*** (4.4E − 05)	0.0004*** (4.6E − 05)	0.0004*** (4.7E − 05)	0.0003*** (9.3E − 05)	0.0003*** (4.7E − 05)
$Log(KSTCK_{t-2})$				0.01758 (0.01689)			
PUBCORE					0.2247** (0.1032)	0.3293*** (0.0977)	0.2284** (0.1034)
$Log(NORM CIT_{t-2})$							−0.1017 (0.1522)
R^2-adj	0.76	0.79	0.79	0.79	0.79		0.79
Log likelihood	−707.30	−670.05	−668.70	−669.51	−667.89		−669.29
Sq. corr.						0.79	
Multicollinearity condition number	7	22	23	27	24		25
F on pooling (time)	0.6694	0.9269	0.6712	0.7141	0.7055		0.4351
F on slope homogeneity	0.2059	0.357	0.2752	0.2683	0.2501		0.2191
White test for heteroscedasticity	44.575***	77.378***	84.013***	92.231***	86.884***		91.191***
LM-Err							
Neigh	0.7199	0.7727	0.7518	0.9808	0.5749		
INV1	3.3586*	2.5407	1.8767	3.4006*	2.6595		
INV2	0.3687	0.9367	0.8782	1.2604	1.020		
LM-Lag							
Neigh	12.214***	3.0067*	2.4689	4.2311**	3.7861*		
INV1	1.6479	0.0642	0.4640	0.061	0.0069		
INV2	5.2928**	0.6649	0.1242	1.9522	1.1352		

Estimated SEs 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 Model 5, the Durbin–Wu–Hausman test for $Log(GRD_{t-2})$ and $Log(GRD_{t-2}) \times NETGRD_{t-2}$ rejects exogeneity at the level of $p < 0.1$. In Model 6, the instruments were selected following the three-group method.

that exogeneity is not rejected by the Durbin–Wu–Hausman test for the variables Log(GRD_{t-2}) and Log(GRD_{t-2}) $\times$ WFP5_Log(RD_{t-2}) the final model 6 is run with 2SLS with heteroscedasticity robust error terms. The dummy variable PUBCORE (with 1 for regions with >1 SD above the EU average publications output and 0 otherwise) enters the equation with significant coefficients in models 5 and 6 suggesting remarkable differences between high and low publishing regions in Europe. All the substantive relationships are confirmed.

With the estimation of model 5a, our intention was to explicitly control for publication quality differences across nations by the inclusion of the variable NORMCIT (normalized number of citations per published paper) into model 5. NORMCIT enters the equation with an insignificant coefficient and statistical properties (in terms of regression fit, parameter sizes, parameter significances) are not affected meaningfully compared to model 5. Thus, our final regression model is robust to cross-national publication quality differences.[16]

The function of Equations (3.5) and (3.6) in the model is to empirically test if regional research productivity exerts long-run cumulative effects on regional knowledge production by attracting further R&D resources directly (Equation 3.5) and pulling knowledge intensive business activities (Equation 3.6) indirectly into the region. Though the empirical literature on the location of the high-technology industry and R&D laboratories has not distilled a 'mainstream' modelling approach (Varga, 2002), a review of recent studies suggests the prominence of three additional classes of factors besides the importance of the regional knowledge base (e.g. R&D intensity/productivity, public research): (i) demand conditions (e.g. market potential: Andersson et al., 2006; Thursby and Thursby, 2006; Siedschlag et al., 2009); (ii) cost conditions (e.g. wages, land prices: Mariani, 2002; Ke and Lai, 2011) and (iii) agglomeration (population density, employment: Ouwersloot and Rietveld, 2000; Cantwell and Piscitello, 2004; Woodward et al., 2006; Borghi et al., 2010). Thus, in our empirically estimated equations in Tables 4 and 5, the additional controls include market potential (MARKETPOT), wages (WAGES) and population density (POPDENS).

In Table 4, we test the effect of R&D productivity on the temporal change of regional R&D expenditures (Equation 3.5). The equation with the changes in R&D expenditures from 1998 to 2001 shows the highest fit and so we report the results for this setup here. The results confirm that the spatial allocation of R&D expenditures is conditioned by R&D productivity, both technological (ALPHAPAT) and scientific (ALPHAPUB). This supports our hypothesized cumulative agglomeration effect behind the temporal changes in regional R&D expenditures. The dummy variable RDHCORE (with 1 for regions with >2 SDs above the EU average R&D expenditures and 0 otherwise) enters the equation with significant coefficients in models 3 and 4, suggesting remarkable differences between high and low R&D performing regions in Europe. We could take this result as an indication of a 'spatial regime effect' favouring high R&D activity regions in the temporal distribution of additional research expenditures. Parameters of the three additional controls are not significant (or only marginally) and do not meaningfully improve regression fit. Thus, we included model 3a only to demonstrate

16 We consider normalized citation per paper as the most appropriate available proxy for publication quality. However an additional regression was also run with citations per paper as an alternative control variable. Regression results are qualitatively the same providing further evidence for model robustness. Regression output is available upon request.

Table 4. Regression results for (GRD2001-GRD1998) for EU regions ($n = 189$)

Model	1	2	3	4	3a
Estimation	OLS	OLS	OLS	OLS-heteroscedasticity robust (white)	OLS (robustness check)
Constant	−604.429*** (90.8252)	−735.41*** (101.405)	−299.107*** (78.3494)	−299.107*** (68.7176)	−299.671*** (81.392)
ALPHAPAT1998	1145.6*** (147.511)	910.258*** (167.819)	351.824*** (125.294)	351.824*** (118.165)	294.301** (140.494)
ALPHAPUB1998		364.853*** (131.181)	190.322** (93.4943)	190.322*** (69.8948)	190.441** (94.147)
RDHCORE			360.98*** (26.3212)	360.98*** (47.4151)	365.808*** (26.451)
MARKETPOT					−2.456 (1.936)
WAGE					2.404* (1.221)
POPDENS					−0.013 (0.013)
R^2-adj	0.24	0.27	0.63	0.63	0.64
White test for heteroscedasticity	52.3206***	57.8899***	42.2263***		87.582***
LM-Err					
Neigh	0.1133	0.0231	0.0674		0.550
INV1	0.0092	0.1976	1.1476		1.552
INV2	0.0895	1.8205	0.9415		1.319
LM-Lag					
Neigh	0.0960	0.0434	0.1026		0.006
INV1	2.6971	0.9635	1.9972		2.620
INV2	0.5956	0.5309	1.9896		1.081

Estimated SEs 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 $p<0.1$.

Table 5. Regression results for (EMPKI2001-EMPKI998) for EU regions ($n = 189$)

Model	1	2	3	4	5
Estimation	OLS	OLS	OLS	ML—spatial error (INV2) with heteroscedasticity weights	ML—spatial error (INV2) with heteroscedasticity weights
Constant	5399.78* (3032.6)	8821.36*** (3314.62)	9955.96*** (3267.78)	11,168.3*** (2879.48)	−3786 (5828.3)
EMPKI1998	0.071*** (0.006)	0.054*** (0.009)	0.032*** (0.012)	0.0262** (0.011)	0.0298*** (0.012)
EMPKI1998 × GRD1998		3.788E − 06** (1.582E − 06)	5.043E − 06*** (1.604E − 06)	5.624E − 06*** (1.604E − 06)	5.327E − 06*** (1.548E − 06)
RDCORE			19,896.5*** (6614.64)	21,321.1*** (6366.96)	14,935.3** (6606.75)
MARKETPOT					−0.008* (0.004)
WAGE					938.004*** (304.377)
LAMBDA				−0.0181** (0.009)	−0.021*** (0.008)
R^2-adj	0.41	0.42	0.45	0.45	0.48
Multicollinearity Condition number	2	4	6	6	9
White test for heteroscedasticity	27.37***	28.182***	34.522***		
LM-Err					
Neigh	0.922	0.164	0.042		
INV1	0.052	0.023	0.28		
INV2	1.008	3.263*	5.878**		
LM-Lag					
Neigh	2.181	1.846	1.916		
INV1	0.479	0.043	0.645		
INV2	4.000*	4.574**	4.316**		

Estimated SEs 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; ***indicates significance at $p < 0.01$; **indicates significance at $p < 0.05$; *indicates $p < 0.1$.

the robustness of model 3. It is noteworthy that spatial dependence is not an issue in any of the models in Table 3, suggesting that the relationship is localized within the boundaries of the region.

In Table 5, we present our estimated model for temporal change in the agglomeration of innovation actors measured by knowledge intensive employment (Equation 3.6). It is clear that strong path dependence is at work in the dynamic distribution of knowledge intensive employment. However, besides this path dependence, the size of regional R&D is also a determining factor as to the direction where knowledge-intensive employment agglomerates. Similar to the results in Table 5, regions with above average R&D expenditures (RDCORE) follow a different pattern in attracting knowledge intensive employment. Both spatial dependence and heteroscedasticity are present consistently throughout models 1–3, which are corrected in the spatial error heteroscedasticity-robust estimation of model 4. Model 5 includes the additional controls with significant parameters. Results suggest that knowledge intensive jobs move to high-wage regions but not necessarily to places with large market potential. This latter result might reflect the fact that technology intensive industries sell their products all over the world and their location decision does not seem to be affected by the level of demand in geographically proximate regions.

6. Simulation analysis: static and dynamic agglomeration and interregional scientific network effects on R&D productivity

The empirical findings so far suggest that regional productivity in Edison-type research (patenting) is influenced by agglomeration but not by interregional scientific networking, whereas regional productivity in Pasteur-type research is influenced by interregional scientific networking but not by agglomeration. How strong are the agglomeration and network effects in each individual region in Europe? Which regions are leading and which ones are lagging behind? On the basis of the above models, we estimated the annual average regional productivity of research in innovation and scientific output for each region using the following formulas:

$$\mathrm{ALPHAPAT}_i = 1.164 \times [0.7088 + 0.1439 \times \mathrm{Log}(\delta_{i,t-2})]^{17}$$

$$\mathrm{ALPHAPUB}_i = [0.4317 + 0.0003 \times \mathrm{NETGRD}_i, t - 2]$$

Our estimates are depicted in the two maps, expressed in SDs from the European mean (Figures 1 and 2). R&D productivity in Edison-type research is more concentrated spatially with core regions in South-West Germany, North-Western Europe (including the South of UK) and the capital city regions. R&D productivity in Pasteur-type research spreads more evenly with less clear spatial concentration patterns indicating that connectedness into interregional scientific networks increases research efficiency in publications even if agglomeration of innovative activities is at a low level.

17 The estimated parameters in Table 2 are multiplied with 1.164. This term is called ‘spatial multiplier’ (Anselin, 2003). It reflects the interdependence among regions in patenting. Interdependence decreases with distance as represented by the squared inverse distance weights matrix in Table 2. Thus, patenting activity is influenced not only by R&D in the region but also by R&D carried out in other regions in the sample following a distance decay pattern.

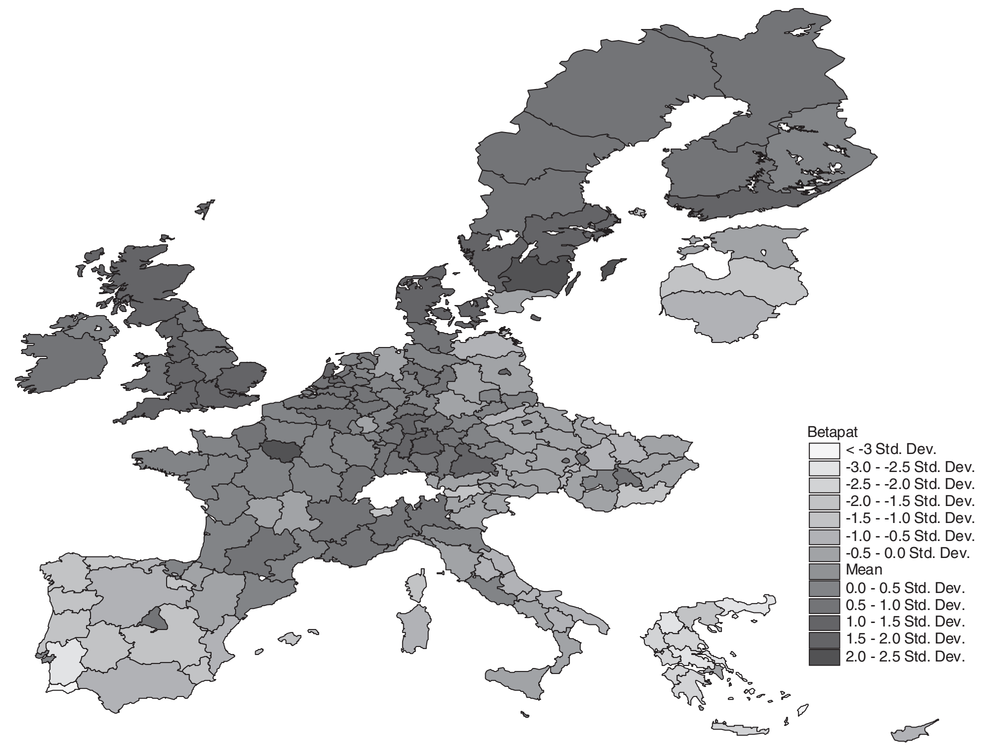

Figure 1. Regional productivity in Edison-type research (patenting).

Importantly, capital cities in East-Central and South Europe are also among the above average R&D productivity regions both in patenting and in publication.

Equations (3.2–3.6) with estimated parameters in Tables 2–5 reflect the dynamic nature of the impacts of R&D support policies. In a relatively short run this support affects patenting directly, while in the longer run it also strengthens concentration of research and knowledge-intensive employment in the region which further impacts knowledge production indirectly (via additional R&D and increased values of the parameters ALPHAPAT and ALPHAPUB). This dynamic feature is represented in Figure 3 where the first seven time periods are shown (without continuing the impacts throughout additional periods).

The econometric estimates allow us to explore counterfactual scenarios and characterize the effects of policy interventions. We produce a simulation of the likely impact of FP6 (2002–2006[18]) funding on patent applications of European regions using the empirically verified relationships and estimated coefficients. We split European regions into four tiers according to their scores on the agglomeration index (δ). Regions with values of the agglomeration index of >1 SD above the mean belong to the first tier. Second tier regions exhibit agglomeration values between the mean and the mean plus 1

18 This is lagged by 1 year (i.e. 2003–2007) in the simulations, better reflecting the period during which the bulk of the funds was spent.

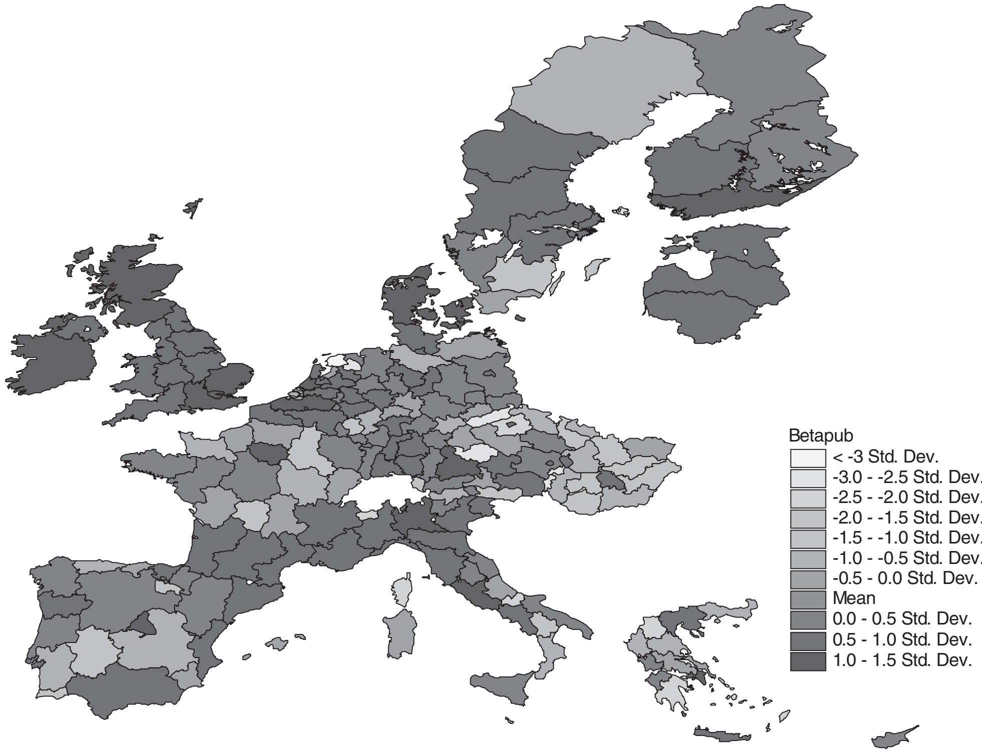

Figure 2. Regional productivity in Pasteur-type research (publications).

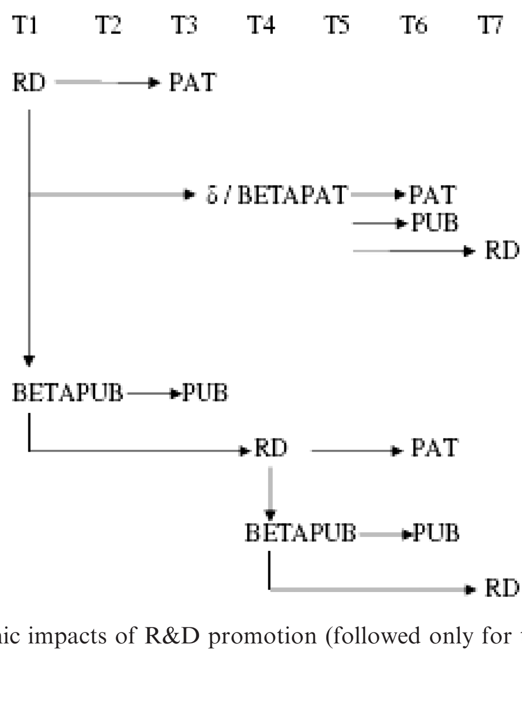

Figure 3. The dynamic impacts of R&D promotion (followed only for the first seven periods).

SD. Third tier regions are half SD value below the mean whereas the rest of the regions belong to the fourth tier.

How effective are European regions in utilizing R&D subsidies awarded from the EU Framework Programs in patenting? Are there differences across regions? How persistent are the impacts over time? The estimated system of equations allows us to calculate a measure of the productivity of FP6 research support in patent applications for each tier and for each year of intervention (2003–2007) and beyond. Simulation results are depicted in Figure 4. Regional productivity of FP6 in patenting is measured by the elasticity of patents with respect to FP6 R&D subsidies.[19]

It is clear from Figure 4 that there are differences across EU regions in the effectiveness of utilizing FP6 R&D subsidies in patenting. Although these differences are relatively minor in the period of intervention (2003–2007), differences in the persistency of the effects are rather significant. Whereas in Tier2 to Tier4 regions the impact of FP6 R&D subsidies on patenting fades away slowly after 2008, Tier1 regions exhibit a persistent (even slightly increasing) impact on patenting. It is the differences in the strengths of the dynamic agglomeration forces that explain the differences in the effectiveness of absorbing R&D subsidies. Whereas Tier1 regions are strong enough to attract additional R&D and human capital that allows them to increase the impact of subsidies on patenting, agglomeration forces in the rest of the regions are not sufficient to maintain even the initial impacts over time.

7. Summary and policy discussion

This article has examined empirically the relative influence of agglomeration and scientific networking on regional R&D productivity in the European Union. The typical data constraints have been tackled by developing and calculating original indices of regional agglomeration of knowledge-producing capabilities using employment data and of interregional networking in R&D using data on R&D collaborations under FP5. The empirical estimation of a system of equations first proposed in Varga (2006) has shed light on three major areas of interest: the relationship between regional agglomeration and interregional scientific networking on the one hand and R&D productivity on the other; the relationship between R&D productivity and temporal changes in regional R&D expenditures; the relationship between R&D expenditures and the generation of knowledge-intensive employment. More specifically, we have estimated KPFs across a number of European regions >3 years testing the influence of agglomeration and scientific networking on the production of Edison- and Pasteur-type knowledge. We found that agglomeration is an important predictor of R&D productivity in the case of Edison-type research, while interregional scientific networking is an important determinant of R&D productivity in the case of Pasteur-type research. Importantly, the two determinants were never *jointly* significant (i.e. interregional scientific networking and agglomeration were not statistically significant for Edison- and Pasteur-type research, respectively)—a finding that is

19 Regional productivity of FP6 R&D support in patenting = [(Estimated number of regional patent applications with FP6—Estimated number of regional patent applications without FP6)/Estimated number of regional patent applications without FP6]/[(Estimated value of regional R&D expenditures with FP6—Estimated value of regional R&D expenditures without FP6)/Estimated number of regional R&D expenditures without FP6].

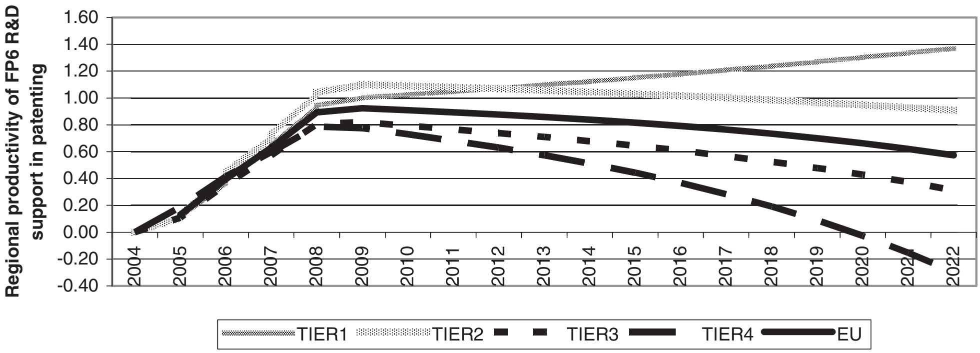

Figure 4. Dynamic agglomeration effects: regional productivity of FP6 R&D support in Edison-type research (patenting).

robust to numerous equation specifications and the choice of stepwise inclusion. This finding indicates that in a knowledge production context and contrary to what may happen in other areas of economic activity (Johansson and Quigley, 2004), agglomeration and scientific networking are neither substitutes nor complements but operate at distinct parts of the knowledge production process.

The sharp contrast between the worlds of Pasteur and Edison raises additional questions that cannot be fully explored here. One may speculate that the distinction is due to a 'hard' constraint on the codifiability of knowledge (Roberts, 2000) and a 'soft' constraint on the willingness of R&D-performing actors to codify knowledge, given the different 'rules of the game' prevalent in the worlds of Pasteur and Edison. Of course, the importance of co-location for knowledge production activities that are heavily dependent on tacit knowledge is recognized in the literature (Malmberg and Maskell, 1997; Morgan, 2004). In the world of Edison, appropriability concerns and a strategy of selective secrecy may also provide part of the explanation. To contrast with, in the world of Pasteur—characterized by fuller disclosure, *de facto* codifiability and the importance of reputation dynamics—access to (not necessarily local) networks makes an important difference.

Our findings with respect to the importance of spatial dependence are in agreement with the above picture: in common with other studies (Paci and Usai, 2000; Maggioni et al., 2007), we find evidence of strong spatial dependence in the production of Edison-type knowledge. As far as the production of Pasteur-type knowledge is concerned though, spatial dependence is either absent or plays a much weaker role.

The latter finding is potentially interesting for scholars studying the importance of local and global effects in cluster and regional economic development. Much of this literature anticipates that a *coincidence* of local 'buzz' *and* access to global 'pipelines' is an important precondition for innovation and economic growth (Bathelt et al., 2004; Yeung et al., 2006). Our findings confirm the long-held suspicion that the respective demand of various industrial sectors for different types of knowledge can be an important predictor of local versus global needs. The distinction between types of knowledge based on motivation of its generation (curiosity or profit), rather than its diffusion prospects (tacit or codified) open new avenues for empirical inquiry in that area. Our findings also provide an interesting new 'lens', through which to interpret

Giuliani's (2007) findings on the asymmetric flow of knowledge within clusters that appear to contradict prevailing views on the importance of spatial distance.

Moreover, using the same sample of regions, we tested empirically the extent to which differences in regional research productivity gives rise to cumulative agglomeration mechanisms in innovation in Europe. That is, we tested whether research productivity acts as a magnet for further research resources and then the resulting increased research activity acts as a centripetal force working towards further agglomeration of regional innovation systems.

Our findings with respect to the spatial allocation of further R&D resources indicate that it is indeed explained by manifested technological and scientific R&D productivity besides a spatial regime effect whereby regions with levels of R&D expenditure that are significantly higher than the sample average attract more funds. We find no evidence of spatial dependence, perhaps a reflection of the high concentration of R&D inputs.

Finally, our empirical test on the relationship between R&D expenditures and the generation of knowledge intensive employment has identified a strongly path-dependent process at work. Past levels of knowledge-intensive employment explain most of the regional variation over time. R&D expenditures though play an important, albeit minor, role in that relationship, as evidenced by the statistically significant interaction between employment and R&D. A spatial regime is also present, whereby regions with levels of R&D expenditure that are significantly higher than the sample average experience greater increases in knowledge intensive employment.

Our results point to the differential effects of agglomeration and scientific networking on the productivity of Edison- and Pasteur-type research. However, it should be noted here that the proxy for interregional networking used in this study, namely research collaboration networks under the EU Framework Programme, poses certain limits to our analysis in the following respects: FP collaboration networks are policy induced and aimed at 'pre-competitive research' of a more 'public good' nature. Other types of research collaboration networks, e.g. co-patenting or co-publication networks may also be found to affect Edison- or Pasteur-type research productivity in ways not explored in the present study. Therefore, the insignificant scientific network effects on Edison-type research productivity found here do not necessarily preclude that other forms of research collaboration play an important role in technological research. We leave this issue for future investigation in a coming paper.

Taken together, the above findings uncover the principal components of regional knowledge production processes across European regions in a dynamic setting. They therefore allow us to explore counterfactual scenarios and characterize the effects of policy interventions. A simulation of the likely impacts of FP6 funds on R&D productivity demonstrates that the dynamic effect is greater in regions with high agglomeration.

A first direct policy conclusion can be derived from the marked differences observed in the production of Edison- and Pasteur-type knowledge. Such differences suggest that a single-pronged instrument that does not distinguish between the two will miss part of its target. The increases in interregional networking that are promoted by the FP appear to have a substantial effect on the productivity of scientific research, but more will need to be done to promote technological research.

Second, the geographical concentration of resources for pre-competitive, Pasteur-type research is at best irrelevant for the generation of new scientific knowledge: in the complex European knowledge production landscape, regions potentially contribute to the creation of scientific knowledge irrespective of their degree of agglomeration. On the other hand, direct funding for competitive, Edison-type research, which from a certain perspective can be seen as a form of hidden industrial subsidy not particularly favoured by the EU competition rules, could only come in practice from national sources, if at all, within the frame of a national innovation and industrial policy. Whether directing funds for Edison-type research at the European level towards highly agglomerated knowledge hubs is an efficient policy option, is a question open to further investigation. As we noted earlier, this question cannot be conclusively answered with the type of research collaboration network used in this study.

A third policy conclusion is drawn from the results of the simulations, which show that the positive effects of collaborative funding instruments, such as the FP, are sustained longer in regions with already high levels of human capital: this indicates that additional attention should be paid to less-advanced regions with the provision of 'structural' funding complementary to the FP, which will be intended to increase the accumulation of human capital and the knowledge capacities of the regions.

Supporting the development of regional innovation capacities in lagging regions will not be easy: our study suggests that regional innovation capacity takes time to develop and comprises a cognitive (knowledge stocks) element as well as an economic element (knowledge-intensive employment). The involvement of diverse policy domains (education, industrial, labour, fiscal policy), the constructive deployment of complementary instruments (direct funding, fiscal incentives, awareness raising) and an intensified coordination of interventions at various levels (European, national, regional) seem necessary.

Acknowledgements

Collaborative research leading to this article started when Attila Varga visited the European Commission's Institute for Prospective Technological Studies in Seville in 2008. Part of this work was completed while Dimitrios Pontikakis was at the Institute for Prospective Technological Studies and at the Newcastle Business School, Northumbria University. We wish to thank Péter Járosi (Faculty of Business and Economics, University of Pécs) for programming the dynamic policy simulations in MATLAB. The regionalization of FP6 participations (used in the simulations) was done jointly by one of the authors and Joaquín M. Azagra-Caro (CSIC). We also wish to express our thanks to participants at all the seminars, workshops and conferences where earlier versions of this article were presented (KITE, University of Newcastle; EC-JRC Institute for Prospective Technological Studies, Seville; EC-DG Regional Policy, Brussels; METU TEKPOL, Ankara; II. World Conference of the Spatial Econometrics Association, New York; II. Conference of the Hungarian Society for Economics, Budapest; Regional Studies Association Annual Conference, Leuven; Annual Meetings European Regional Science Association, Łódz; DIME Workshop 'Measuring, mapping and modelling technical change', Sant'Anna School of Advanced Studies, Pisa; KRTI, University of Pécs, Faculty of Business and Economics).

Disclaimer

The views expressed in this paper are purely those of the authors and may not in any circumstances be regarded as stating an official position of the OECD.

Funding

Funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 216813 is gratefully acknowledged.

References

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

Acs, Z., Varga, A. (2005) Entrepreneurship, agglomeration and technological change. *Small Business Economics*, 24: 323–334.

Amin, A., Thrift, N. (1992) Neo-Marshallian nodes in global networks. *International Journal of Urban and Regional Research*, 16: 571–587.

Andergassen, R., Nardini, F., Ricottilli, M. (2005) Firms' network formation through the transmission of heterogeneous knowledge. Working Paper No. 543, Dipartimento di Scienze Economiche, Università degli Studie.

Andersen, E. S. (1992) Approaching the national systems of innovation from the production and linkage structure. In B. Å. Lundvall (ed.) *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. London: Pinter.

Andersson, M., Gråsjö, U., Karlsson, C. (2006) Industry R&D Location—the role of accessibility to university R&D and institutions of higher education CESIS Electronic Working Paper Series, Paper No.68.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Anselin, L. (2003) Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26: 153–166.

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

Audretsch, D., Feldman, M. (1996) R&D spillovers and the geography of innovation and production. *American Economic Review*, 86: 630–640.

Autant-Bernard, C., Billand, P., Frachisse, D., Massard, N. (2007) Social distance versus spatial distance in R&D cooperation: empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in Regional Science*, 86: 495–519.

Balconi, M., Breschi, S., Lissoni, F. (2004) Networks of inventors and the role of academia: an exploration of Italian patent data. *Research Policy*, 33: 127–145.

Barabási, A., Albert, R. (1999) Emergence of scaling in random networks. *Science*, 286: 509–512.

Barber, M. J., Krueger, A., Krueger, T., Roediger-Schluga, T. (2006) Network of European Union-funded collaborative research and development projects. *Physical Review E*, 73: 036132/1–13.

Barré, R., Henriques, L., Pontikakis, D., Weber, M. (forthcoming) Measuring the integration and coordination dynamics of the European Research Area. *Science and Public Policy*.

Bathelt, H., Malmberg, A., Maskell, P. (2004) Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28: 31–56.

Belsley, D., Kuh, E., Welsch, R. (1980) *Regression Diagnostics, Identifying Influential Data and Sources of Collinearity*. New York: Willey.

Billand, P., Frachisse, D., Massard, N. (2008) The Sixth Framework Program as an affiliation network: representation and analysis. FEEM Fondazione Eni Enrico Mattei Research Paper Series, No. 32.2008.

Borghi, E., del Bo, C., Florio, M. (2010) Industrial clusters and regional innovation: An evaluation and implications for economic cohesion. Dipartimento Di Scienze Economiche Aziendali E Statistiche, Universitá Degli Studi De Milano, WP 2010-18.

Breschi, S., Lissoni, F. (2005) Mobility and social networks: localised knowledge spillovers revisited. CESPRI Working Paper 142, Università Commerciale 'Luigi Bocconi'.

Caloghirou, Y., Vonortas, N. S., Ioannides, S. (eds) (2004) *European Collaboration in Research and Development*. Northampton: Edward Elgar.

Cantwell, J., Piscitello, L. (2005) Recent location of foreign-owned research and development activities by large multinational corporations in the european regions: the role of spillovers and externalities. *Regional Studies*, 39: 1–16.

Carayol, N., Roux, P. (2007) The strategic formation of inter-individual collaboration networks. Evidence from co-invention patterns. *Annales d'Économie et de Statistique*, 87(88): 275–301.

Chorafakis, G., Pontikakis, D. (2011) Theoretical underpinnings and future directions of EU research policy: a paradigm shift? *Prometheus*, 29: 131–161.

Cooke, P. (2001) Regional innovation systems, clusters and the knowledge economy. *Industrial and Corporate Change*, 10: 945–974.

Cowan, R. (2004) Network models of innovation and knowledge diffusion. Maastricht Economic Research Institute on Innovation and Technology (MERIT). Research Memoranda Series, No. 16.

Cowan, R., Jonard, N. (2006) Innovation networks and the distribution of knowledge. Paper presented at DRUID Summer Conference on: *Knowledge, Innovation and Competitiveness: Dynamics of Firms, Networks, Regions and Institutions*. Denmark: Copenhagen. June 18–20. Available online at: http://www2.druid.dk/conferences/viewpaper.php?id=145&cf=8.

Crescenzi, R., Rodríguez-Pose, A., Storper, M. (2007) The territorial dynamics of innovation: a Europe-United States comparative analysis. *Journal of Economic Geography*, 7: 673–709.

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

Dosi, G. (1988) Sources, procedures, and microeconomic effects of innovation. *Journal of Economic Literature*, 26: 1120–1171.

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

Ellison, G., Glaeser, E. (1997) Geographic concentration in US manufacturing industries: a Dartboard approach. *Journal of Political Economy*, 105: 889–927.

European Commission (2009) A more research-intensive and integrated European Research Area, Science, Technology and Competitiveness key figures report 2008/2009. EUR 23608, Luxembourg, Office for Official Publications of the European Communities.

European Commission (2010a) Europe 2020 Flagship Initiative Innovation Union SEC(2010)1161, European Commission, Brussels.

European Commission (2010b) Regional Policy Contributing to Smart Growth in Europe 2020. SEC(2010)1183, European Commission, Brussels.

European Commission (2011) Innovation Union Competitiveness Report, Publication Office of the European UnionLuxembourg.

Fafchamps, M., van der Leij, M., Goyal, S. (2006) Scientific networks and co-authorship. University of Oxford, Department of Economics, Economics Series Working Papers, No. 56.

Feldman, M. (1994) The university and economic development: the case of Johns Hopkins University and Baltimore. *Economic Development Quarterly*, 8: 67–66.

Foray, D. (2009) Understanding Smart Specialisation. In Pontikakis, Kyriakou, van Bavel (eds) (2009): The question of R&D specialisation: perspectives and policy implications, JRC Scientific and Technical Reports, No. 51665, EUR 23834, Luxembourg: Office for Official Publications of the European Communities. Available online at: http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2339.

Foray, D., 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. Available online at: http://ec.europa.eu/invest-in-research/pdf/download_en/policy_brief1.pdf.

Fujita, M., Krugman, P., Venables, A. (1999) *The Spatial Economy*. Cambridge, London: MIT Press.

Fujita, M., Thisse, J. (2002) *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth*. Cambridge, MA, London, England: Cambridge University Press.

Furman, J. L., Hayes, R. (2004) Catching up or standing still? National R&D productivity among 'follower' countries, 1978–1999. *Research Policy*, 33: 1329–1354.

Furman, J. L., Porter, M. E., Stern, S. (2002) The determinants of national innovative capacity. *Research Policy*, 31: 899–933.

Gastner, M. T., Newman, M. E. (2006) The spatial structure of networks. *European Physical Journal B*, 49: 247–252.

Georghiou, L., Cooke, P., Cozzens, S., Dearing, A., Henriques, L., Langer, J., Laredo, P., Sanz Menendez, L., Weber, M. (2008) Challenging Europe's research: rationales for the European Research Area (ERA) Report of the ERA, Expert Group, Available online at: http://ec.europa.eu/research/era/pdf/eg7-era-rationales-final-report_en.pdf.
Geroski, P. A. (1989a) European industrial policy and industrial policy in Europe. *Oxford Review of Economic Policy*, 5: 20–36.
Geroski, P. A. (1989b) The choice between diversity and scale. In E. Davis, P. A. Geroski, J. A. Kay, A. Manning, C. Smales, S. R. Smith, S. Szymanski (eds) *Myths and Realities, Centre for Business Strategy*, pp. 29–45. London: London Business School.
Giuliani, E. (2007) The selective nature of knowledge networks in clusters: evidence from the wine industry. *Journal of Economic Geography*, 7: 139–168.
Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., Shleifer, A. (1992) Growth in cities. *Journal of Political Economy*, 100: 1126–1152.
Goldstein, H., Drucker, J. (2006) The economic development impacts of universities on regions: do size and distance matter? *Economic Development Quarterly*, 20: 22–43.
Goyal, S., Moraga-Gonzalez, J. L. (2001) R&D networks. *RAND Journal of Economics*, 32: 686–707.
Granovetter, M. S. (1973) The strength of weak ties. *American Journal of Sociology*, 78: 1360–1380.
Griliches, Z. (1990) Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, 28: 1661–1707.
Hall, B. H. (1993) The Stock Market's Valuation of R&D Investment During the 1980's. *The American Economic Review*, 83: 259–264.
Hall, B. H., Griliches, Z., Hausman, J. A. (1986) Patents and R and D: is there a lag? *International Economic Review*, 27: 265–283.
Hite, J. M., Hesterly, W. S. (2001) The evolution of firm networks: from emergence to early growth of the firm. *Strategic Management Journal*, 22: 275–286.
Jacobs, J. (1969) *The Economy of Cities*. New York: Random House.
Jaffe, A. (1989) Real effects of academic research. *American Economic Review*, 79: 957–970.
Jaffe, A., Trajtenberg, M., Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, 108: 577–598.
Johansson, B., Quigley, J. M. (2004) Agglomeration and networks in spatial economies. *Papers in Regional Science*, 83: 165–176.
Jones, C. (1995) R&D-based models of economic growth. *Journal of Political Economy*, 103: 759–784.
Ke, S., Lai, M. (2011) Productivity of Chinese regions and the location of multinational research and development. *International Regional Science Review*, 34: 102–131.
Kennedy, P. (1998) *A Guide to Econometrics*. Malden: Blackwell Publishing.
Koo, J. (2005) Agglomeration and spillovers in a simultaneous framework. *Annals of Regional Science*, 39: 35–47.
Luukkonen, T., Persson, O., Sivertsen, G. (1992) Understanding patterns of international scientific collaboration. *Science, Technology and Human Values*, 17: 101–126.
Maggioni, M. A., Nosvelli, M., Uberti, T. E. (2007) Space versus networks in the geography of innovation: a European analysis. *Papers in Regional Science*, 86: 471–493.
Malmberg, A., Maskell, P. (1997) Towards an explanation of regional specialization and industry agglomeration. *European Planning Studies*, 5: 25–41.
Mariani, M. (2002) Next to Production or to Technological Clusters? The Economics and Management of R&D Location. *Journal of Management and Governance*, 6: 131–152.
Markusen, A. (1996) Sticky places in slippery space: a typology of industrial districts. *Economic Geography*, 72: 293–313.
Marshall, A. (1920) *Principles of economics*. London: Macmillan.
Matthews, M., McGowan, F. (1992) Reconciling diversity and scale: some questions of method in the assessment of the costs and benefits of European integration. *Revue d'économie industrielle*, 59: 222–234.
May, R. M. (1997) The scientific wealth of nations. *Science*, 275: 793–796.
McCann, P., Ortega-Argilés, R. (2011) Smart specialisation, regional growth and applications to EU Cohesion Policy. Economic Geography Working Paper 2011, Faculty of Spatial Sciences, University of Groningen.

Moed, H. F., De Bruin, R. E., Van Leeuwen, T. N. (1995) New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications. *Scientometrics*, 33: 381–422.
Mokyr, J. (2002) *The Gifts of Athena*. Princeton: Princeton University Press.
Morgan, K. (2004) The exaggerated death of geography: learning, proximity and territorial innovation systems. *Journal of Economic Geography*, 4: 3–21.
Nelson, R. R. (ed.) (1993) *National Innovation Systems: A Comparative Analysis*. Oxford: Oxford University Press.
Newman, M. E. J. (2000) Models of the small world: a review. *Journal of Statistical Physics*, 101: 819–841.
Newman, M. E. J. (2001) The structure of scientific collaboration networks. *PNAS*, 98: 404–409.
OECD (2012) *OECD Science, Technology and Industry Scoreboard* 2011. *Innovation and Growth in Knowledge Economies*. Paris: OECD Publications.
Ouwersloot, H., Rietweld, P. (2000) The geography of R&D. Tobit analysis and Bayesian approach to mapping R&D activities for The Netherlands. Tinbergen Institute Discussion Paper, TI 2000-043/3.
Paci, R., Usai, S. (2000) Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal*, 49: 381–390.
Park, G., Park, Y. (2006) On the measurement of patent stock as knowledge indicators. *Technological Forecasting and Social Change*, 73: 793–812.
Ponds, R., van Oort, F., Frenken, K. (2007) The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86: 423–443.
Ponds, R., van Oort, F. G., Frenken, K. (2009) Internationalization and regional embedding of scientific research in the Netherlands. In A. Varga (ed.) *Universities and Regional Development*, pp. 109–137. Cheltenham: Edward Elgar.
The question of R&D specialisation: perspectives and policy implications. Pontikakis, D., Kyriakou, D., van Bavel, R. (eds) (2009) *JRC Scientific and Technical Reports*. No. 51665, EUR 23834, Luxembourg: Office for Official Publications of the European Communities. Available online at: http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2339.
Porter, M. E., Stern, S. (2000) Measuring the 'ideas' production function: evidence from international patent output. NBER Working Papers, No. 7891, National Bureau of Economic Research.
Roberts, J. (2000) From know-how to show-how? Questioning the role of information and communication technologies in knowledge transfer. *Technology Analysis and Strategic Management*, 12: 429–443.
Romer, P. M. (1990) Endogenous technological change. *Journal of Political Economy*, 5: S71–S102.
Rosenthal, S., Strange, W. (2004) Evidence on the nature and sources of agglomeration economies. In V. Henderson, J. Thisse (eds) *Handbook of Regional and Urban Economics*, 4: 2119–2171.
SCImago. (2012) SJR—SCImago journal & country rank. In I. Siedschlag, D. Smith, C. Turcub and Zhanga, X. (eds) *What Determines the Attractiveness of the European Union to the Location of R&D Multinational Firms?* ESRI Working Paper No. 306. Available at: http://www.scimagojr.com [Accessed 17 August].
Siedschlag, I., Smith, D., Turcub, C., Zhanga, X. (2009) What Determines the Attractiveness of the European Union to the Location of R&D Multinational Firms? ESRI Working Paper No. 306.
Stokes, D. E. (1997) *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington: Brookings Institution Press.
Storper, M. (1997) *The Regional World*. London: Guilford Press.
Stuart, T. E., Sorenson, O. (2007) Strategic networks and entrepreneurial ventures. *Strategic Entrepreneurship Journal*, 1: 211–227.
Taylor, M. Z. (2009) International linkages and national innovation rates: an exploratory probe. *Review of Policy Research*, 26: 127–149.
Thursby, J., Thursby, M. (2006) *Here or There? A survey of factors in multinational R&D location*. Report to the Government-university-industry research roundtable. Washington, D.C: The National Academies Press.

van Raan, A. F. J. (2004) Measuring science. In H. F. Moed, W. Glänzel, U. Schmoch (eds) *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*, pp. 19–50. Dordrecht: Kluwer Academic Publishers.
Varga, A. (1999) Time-space patterns of US innovation: stability or change? A detailed analysis based on patent data. In M. Fischer, L. Suarez-Villa, M. Steiner (eds) *Innovation, Networks and Localities*, pp. 215–234. Berlin: Springer.
Varga, A. (2000) Local academic knowledge transfers and the concentration of economic activity. *Journal of Regional Science*, 40: 289–309.
Varga, A. (2001) Universities and regional economic development: does agglomeration matter? In B. Johansson, C. Karlsson, R. Stough (eds) *Theories of Endogenous Regional Growth—Lessons for Regional Policies*, pp 345–367. Berlin: Springer.
Varga, A. (2002) Knowledge transfers from universities to the regional economy: a review of the literature. In A. Varga, L. Szerb (eds) *Innovation, Entrepreneurship and Regional Economic Development: International Experiences and Hungarian Challenges*, pp. 147–171. Pécs: University of Pécs Press.
Varga, A. (2006) The spatial dimension of innovation and growth: empirical research methodology and policy analysis. *European Planning Studies*, 9: 1171–1186.
Varga, A., Anselin, L., Acs, Z. (2005) Regional innovation in the US over space and time. In G. Maier, S. Sedlacek (eds) *Spillovers and Innovation: City, Environment, and the Economy. Interdisciplinary Studies in Economics and Management*, pp. 93–104. Wien, New York: Springer.
Varga, A., Parag, A. (2009) Academic knowledge transfers and the structure of international research networks. In A. Varga (ed.) *Universities and Regional Development*, pp. 138–159. Cheltenham: Edward Elgar.
Varsakelis, N. C. (2006) Education, political institutions and innovative activity: a cross-country empirical investigation. *Research Policy*, 35: 1083–1090.
Wagner, C. S., Leydesdorff, L. (2005) Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, 34: 1608–1618.
White, H. C. (1992) *Identity and Control*. Princeton: Princeton University Press.
Woodward, D., Figueiredo, O., Gumaraes, P. (2006) Beyond the Silicon Valley: University R&D and high-technology location. *Journal of Urban Economics*, 60: 15–32.
Yeung, H. W., Liu, W., Dicken, P. (2006) Transnational corporations and network effects of a local manufacturing cluster in mobile telecommunications equipment in China. *World Development*, 34: 520–540.
Zucker, L. G., Darby, M. R., Furner, J., Liu, R. C., Ma, H. (2007) Minerva unbound: knowledge stocks, knowledge flows and new knowledge production. *Research Policy*, 36: 850–863.

Appendix

Table A1. Descriptive statistics

Variable	Mean	SD (overall)	SD (between)	SD (within)	Min.	Max.	*n*
PAT	318.4363	536.1444	535.7374	38.08507	0.01	3460.89	567
PUB	1921.995	2531.388	2528.203	196.7256	1	22,022	567
GRD	693.127	1169.854	1170.091	65.43073	1	11,436	567
PSTCK	27,429.94	33,173.6	33,045.87	3509.518	6	98,481	567
EMPKI	346,197.6	364,772.9	365,110.3	14,992.06	2696	25,52324	567
Δ	0.968575	0.293157	0.291911	0.032097	0.275	1.982	567
NETGRD	781.1124	229.4251	229.724	7.015673	55.167	1045.984	567
PATCORE	0.275132	0.446975			0	1	567
RDCORE	0.291005	0.454627			0	1	567
PUBCORE	0.349206	0.47714			0	1	567
PATHCORE	0.10582	0.307879			0	1	567
RDHCORE	0.10582	0.307879			0	1	567
PUBHCORE	0.10582	0.307879			0	1	567
ALPHAPAT98	0.649286	0.034698			0.52492	0.72627	189
ALPHAPUB98	0.754445	0.090864			0.46707	0.85308	189

Table A2. List of regions

NUTS code	Region
AT11	Burgenland
AT12	Niederösterreich
AT13	Wien
AT21	Kärnten
AT22	Steiermark
AT31	Oberösterreich
AT32	Salzburg
AT33	Tirol
AT34	Vorarlberg
BE1	Région de Bruxelles-Capitale
BE2	Prov. Antwerpen
BE3	Prov. Brabant Wallon
CY00	Kypros/Kibris
CZ01	Praha
CZ02	Střední Čechy
CZ03	Jihozápad
CZ04	Severozápad
CZ05	Severovýchod
CZ06	Jihovýchod
CZ07	Střední Morava
CZ08	Moravskoslezsko
DE11	Stuttgart
DE12	Karlsruhe
DE13	Freiburg
DE14	Tübingen
DE21	Oberbayern
DE22	Niederbayern
DE23	Oberpfalz
DE24	Oberfranken
DE25	Mittelfranken
DE26	Unterfranken
DE27	Schwaben
DE30	Berlin
DE4	Brandenburg
DE50	Bremen
DE60	Hamburg
DE71	Darmstadt
DE72	Gießen
DE73	Kassel
DE80	Mecklenburg-Vorpommern
DE91	Braunschweig
DE92	Hannover
DE93	Lüneburg
DE94	Weser-Ems
DEA1	Düsseldorf
DEA2	Köln
DEA3	Münster
DEA4	Detmold
DEA5	Arnsberg
DEB1	Koblenz
DEB2	Trier

(continued)

Table A2. Continued

NUTS code	Region
DEB3	Rheinhessen-Pfalz
DEC0	Saarland
DED1	Chemnitz
DED2	Dresden
DED3	Leipzig
DEE	Sachsen-Anhalt
DEF0	Schleswig-Holstein
DEG0	Thüringen
DK00	Danmark
EE00	Eesti
ES11	Galicia
ES12	Principado de Asturias
ES13	Cantabria
ES21	País Vasco
ES22	Comunidad Foral de Navarra
ES23	La Rioja
ES24	Aragón
ES30	Comunidad de Madrid
ES41	Castilla y León
ES42	Castilla-La Mancha
ES43	Extremadura
ES51	Cataluña
ES52	Comunidad Valenciana
ES53	Illes Balears
ES61	Andalucía
ES62	Región de Murcia
FI13	Itä-Suomi
FI18	Etelä-Suomi
FI19	Länsi-Suomi
FI1A	Pohjois-Suomi
FI20	Åland
FR10	Île de France
FR21	Champagne-Ardenne
FR22	Picardie
FR23	Haute-Normandie
FR24	Centre
FR25	Basse-Normandie
FR26	Bourgogne
FR30	Nord—Pas-de-Calais
FR41	Lorraine
FR42	Alsace
FR43	Franche-Comté
FR51	Pays de la Loire
FR52	Bretagne
FR53	Poitou-Charentes
FR61	Aquitaine
FR62	Midi-Pyrénées
FR63	Limousin
FR71	Rhône-Alpes
FR72	Auvergne
FR81	Languedoc-Roussillon

(continued)

Table A2. Continued

NUTS code	Region
FR82	Provence-Alpes-Côte d'Azur
FR83	Corse
GR11	Anatoliki Makedonia, Thraki
GR12	Kentriki Makedonia
GR13	Dytiki Makedonia
GR14	Thessalia
GR21	Ipeiros
GR23	Dytiki Ellada
GR24	Sterea Ellada
GR25	Peloponnisos
GR30	Attiki
GR42	Notio Aigaio
GR43	Kriti
HU10	Közép-Magyarország
HU21	Közép-Dunántúl
HU22	Nyugat-Dunántúl
HU23	Dél-Dunántúl
HU31	Észak-Magyarország
HU32	Észak-Alföld
HU33	Dél-Alföld
IE	Ireland
ITC1	Piemonte
ITC2	Valle d'Aosta/Vallée d'Aoste
ITC3	Liguria
ITC4	Lombardia
ITD1	Provincia Autonoma Bolzano/Bozen
ITD2	Provincia Autonoma Trento
ITD3	Veneto
ITD4	Friuli-Venezia Giulia
ITD5	Emilia-Romagna
ITE1	Toscana
ITE2	Umbria
ITE3	Marche
ITE4	Lazio
ITF1	Abruzzo
ITF2	Molise
ITF3	Campania
ITF4	Puglia
ITF5	Basilicata
ITF6	Calabria
ITG1	Sicilia
ITG2	Sardegna
LT00	Lietuva
LU00	Luxembourg (Grand-Duché)
LV00	Latvija
MT00	Malta
NL11	Groningen
NL12	Friesland
NL13	Drenthe
NL21	Overijssel
NL22	Gelderland

(continued)

Table A2. Continued

NUTS code	Region
NL23	Flevoland
NL31	Utrecht
NL32	Noord-Holland
NL33	Zuid-Holland
NL34	Zeeland
NL41	Noord-Brabant
NL42	Limburg (NL)
PT11	Norte
PT15	Algarve
PT16	Centro (P)
PT17	Lisboa
PT18	Alentejo
SE01	Stockholm
SE02	Östra Mellansverige
SE04	Sydsverige
SE06	Norra Mellansverige
SE07	Mellersta Norrland
SE08	Övre Norrland
SE09	Småland med öarna
SE0A	Västsverige
SK01	Bratislavský kraj
SK02	Západné Slovensko
SK03	Stredné Slovensko
SK04	Východné Slovensko
UKF	Lincolnshire
UKG	Shropshire and Staffordshire
UKH	East Anglia
UKI	Inner London
UKJ	Surrey, East and West Sussex
UKK	Cornwall and Isles of Scilly
UKL	West Wales and The Valleys
UKM	Eastern Scotland
UKN	Northern Ireland
UKC	Northumberland and Tyne and Wear
UKD	Cumbria
UKE	West Yorkshire