Embeddedness of regions in European knowledge networks. A comparative analysis of inter-regional R&D collaborations, co-patents and co-publications

07.13

Iris Wanzenböck, Thomas Scherngell and Thomas Brenner

Philipps Universität Marburg

Impressum:

Working Papers on Innovation and Space
Philipps-Universität Marburg

Herausgeber:

Prof. Dr. Dr. Thomas Brenner
Deutschhausstraße 10
35032 Marburg
E-Mail: thomas.brenner@staff.uni-marburg.de

Erschienen: 2013

Embeddedness of regions in European knowledge networks. A comparative analysis of inter-regional R&D collaborations, co-patents and co-publications

Iris Wanzenböck[1], Thomas Scherngell

both from Foresight and Policy Development Departement, Austrian Institute of Technology (AIT), Vienna.

Thomas Brenner

Economic Geography and Location Research, Philipps-University, Marburg.

Abstract:

This paper investigates the embeddedness of European regions in different types of inter-regional knowledge networks, namely project based R&D collaborations within the EU Framework Programmes (FPs), co-patent networks and co-publication networks. Embeddedness refers to the network positioning of regions captured in terms of social network analytic (SNA) centrality measures. The objective is to estimate how region-internal and region-external factors influence network embeddedness in the distinct network types, in order to identify differences in their driving factors at the regional level. In our modelling approach, we apply advanced spatial econometric techniques by means of a mixed effects panel version of the Spatial Durbin Model (SDM), and introduce a set of variables accounting for a capacity-specific, a relational as well as a spatial dimension in regional knowledge production activities. The results reveal conspicuous differences between the knowledge networks. Internal capacity- and technology-related aspects but also spatial spillover impacts from surrounding regions prove to be particularly important for centrality in the co-patent network. We also find significant - region-internal and region-external - impacts of general economic conditions on a region's centrality in the FP network. However, we cannot observe substantial spillover effects of region-external factors on centrality in the co-publication network. Thus, the distinctive knowledge creation foci in each network seem to find expression in the network structure as well as its regional determinants.

Keywords: knowledge networks, network embeddedness, network centrality, regional knowledge production, panel Spatial Durbin model.

JEL Classifications: L14, N74, O33, R15

---

[1] Corresponding Author: Iris Wanzenböck, Foresight and Policy Development Departement, Austrian Institute of Technology (AIT), Vienna, E-Mail: Iris.Wanzenboeck.fl@ait.ac.at.

1 Introduction

The geography of knowledge networks has attracted much attention in theoretical and empirical scientific works in the recent past (see Scherngell and Barber 2009, Hoekman et al. 2009, Autant-Bernard et al. 2007), related to considerations that networks of research actors are essential for the creation of new knowledge (see, for instance, Powell and Grodal 2005), and thus, fundamental for economic performance of firms as well as regions or countries (Romer 1990). Knowledge networks – defined as a set of actors jointly producing knowledge, for instance in form of joint R&D projects, joint publications or joint applications for patents – may constitute promising vehicles for research actors to tap knowledge that is widely dispersed in geographical space. This knowledge diffuses through networks not only via direct links but also through indirect allies. Thus, the embeddedness and strategic positioning in such knowledge networks is of central importance to enhance benefits of knowledge access.

The focus of this study is on the embeddedness in different types of knowledge networks at the regional level. By taking a regional perspective, we assume that participation of organisations in networks enriches not only firm-internal knowledge creation processes, but has also significant influence on the innovation capacity of entire regions, particularly due to the presence of geographically localised knowledge spillovers (Ashheim et al. 2011, Karlsson and Manduchi 2001, Lagendijk 2001). Collaboration intensive organisations and their positioning in such knowledge networks are important levers for knowledge diffusion in their local, intra-regional environment, i.e knowledge gained by inter-regional network channels is likely to be injected to intra-regional knowledge diffusion mechanisms (see, for example, Breschi and Lissoni 2009, Bathelt et al. 2004, Varga 2001). In this sense, knowledge creation of regions depends not only on internal conditions, but also on the ability of localised actors to identify and quickly access region-external knowledge sources, and thus, on their ability to participate in collaborations as well as their positioning in inter-regional knowledge networks.

The question that arises in this context is which factors affect the positioning of a region – referred to as a region's embeddedness – in such inter-regional knowledge networks. Thus, in this study we aim to identify distinctive driving factors of regional network embeddedness from a comparative perspective on three different inter-regional knowledge networks across Europe. We define inter-regional knowledge networks as a system consisting of a set of regions (vertices in the network) that are connected to each other by collaborative research endeavours (edges in the network) in the form of project based R&D collaborations funded by the EU Framework Programmes (FP), co-patens or co-publications. Network embeddedness is then defined in terms of centrality as used in Social Network Analysis (SNA). In SNA it is assumed that vertices having a central network position will more likely benefit from network advantages than actors that have a more de-central, peripheral position in the network (see, for instance, Wasserman and Faust 1994). Thus, a stronger embeddedness of a region may increase information and knowledge access within the network. Such a privileged position provides opportunity for regions to exploit knowledge flows passing through network ties and to faster realise relevant region-external knowledge sources.

Following conventional regional knowledge production function approaches (Autant-Bernard 2012), we employ a spatial econometric perspective to identify distinctive driving factors of regional network embeddedness, taking account of region-internal capacities and relational factors of knowledge production activities, as well as spatial, region-external characteristics. To distinguish and measure magnitude and significance of the direct and spatial spillover effects over space and time, we employ a mixed effects panel version of the Spatial Durbin Model (SDM) (see Elhorst 2003 and 2012). The dependent variable is measured in terms of a region's centrality in the distinct inter-regional knowledge networks for the years 1999-2006, using a regional setting of 241 NUTS-2 regions of the EU-25 member states. We rely on two different centrality concepts, that are betweenness- and eigenvector centrality (see, Wasserman and Faust 1994), in order to regard network embeddedness from different network theoretical perspective.

The study contributes to existing related literature (see Bergman and Maier 2009, Wanzenböck et al. 2012) by at least three respects: *First*, by focusing on three distinct network types, we are able to investigate the structural differences of different knowledge networks in both relational and spatial terms. *Second*, we analyse these knowledge networks in their European dimension, providing a comprehensive picture on the spatial spread of these networks across Europe, which is not least an important issue in light of current policy endeavours concerning a closer integration of the European science and R&D landscape. In this regard, *third*, the study will enrich our understanding on determinants that discriminate core regions from less dominant regions in Europe, delivering major contributions to the discussion on why certain regions are more efficient in creating new knowledge.

The remainder of the study is organised as follows. Section 2 sets forth the theoretical background, embeds the current study in related literature and derives the main hypotheses for the empirical analysis. Section 3 operationalises the concept of regional embeddedness in knowledge networks. Then, Section 4 deals with the major distinctive characteristics of the three knowledge networks under consideration. Section 5 introduces the data, gives a first descriptive analysis of our network centrality measures and presents exploratory spatial analysis on network centrality of European regions, before Section 6 describes the mixed effects panel version of the Spatial Durbin Model (SDM) and introduces the set of independent variables. Section 7 presents the panel SDM estimation results as well as the associated impact estimates for the three knowledge networks under consideration. Section 8 concludes with the main results from a comparative perspective and ideas for a future research agenda.

2 Regional positioning in knowledge networks

Recent decades have seen a surge of collective efforts in knowledge production that are distributed across organisations and spanning distant locations (Wuchy et al. 2007, Hagedoorn et al. 2000, Hicks and Katz 1996). Scientific, research and innovation activities are increasingly inter-organisational processes among firms, universities and research organisations at the regional, national or even global level. Knowledge networks – manifested in form of collaborations to create new knowledge, for instance, in the form of joint R&D projects, joint assignment of patents or co-publications – connect localized actors to external knowledge bases that are dispersed in geographical space. Thus, such collaborations are regarded to be one of the main carriers of long-distance knowledge (Breschi and Lissoni 2009, Ponds et al. 2007). A portfolio of external relations is widely considered to be a crucial asset in knowledge production processes, particularly due to the opportunity to learn from each other, to pool resources and skills as well as to get access to specific knowledge components in a flexible and purposeful way (Zucker and Darby 2007, Katz and Martin 1997).

Apart from the mentioned advantages of direct and bilateral relations, such knowledge networks provide wider opportunities to stimulate knowledge creation, particularly due to the facilitated access to information and knowledge distributed through network links, also through indirect paths. Network relations accelerate the search for synergies between own intentions and the activities of others and eases selection and formation of new ties (Gilsing et al. 2008, Borgatti and Everett 2006, Sorenson et al. 2006). Likewise, attaching with core players and elite researchers may be more probable due to common acquaintances (in network theory such phenomena are often referred to as preferential attachment or friends-of-friends structures; see, for example, Ter Wal and Boschma 2009). Moreover, participation in networks enables to keep abreast with the state-of-the-art in the research field, as information on scientific and technological advances leaks more likely via network allies (Powell and Giannella 2010, Katz and Martin 1997). Along these channels, own ideas may also be carried forth to a wider community, enhancing an actor's visibility in the network.

From this perspective, actors intensively involved in several collaborative arrangements are directly interlinked with others, show short pathways to diverse sets of network nodes, and therefore, are highly embedded in the network structure. Central players act either as hubs for knowledge diffusion, spreading knowledge throughout several connected actors, or / and are in a position to enable but also control knowledge flows between various de facto unconnected allies, acting as a 'gatekeeper' for information and knowledge running through them. In both positions they exert influence on the process of knowledge transmission throughout the entire network (Borgatti and Everett 2006).

With the above in mind, it seems appropriate to take a network perspective when analysing knowledge diffusion resulting from research collaborations, considering not only direct connections, but also access to 'second-hand' information and knowledge. One crucial question that arises in this context concerns the conditions that drive central positioning in such knowledge networks. In this study, we take a Regional Science perspective and focuse on inter-regional knowledge networks across Europe, composed of research actors located in one region and their region-external knowledge links. We consider region-internal and region-external conditions that drive a region's embeddededness in inter-regional knowledge networks. The following streams in the literature provide the theoretical foundations for our empirical study, leading to three main dimensions determining a region's network embeddedness.

The capacity dimension

Resource-based approaches on research and innovation emphasise that technological capacity and skills of actors are one of the most significant factors to absorb, exploit and assemble different types of knowledge produced by others (Cohen and Levinthal 1990). This involves also the resources and capacity to build up a variation of ties, which are required in order to receive and further transmit knowledge flowing through numerous network channels. We follow such resource-related approaches at the regional level (Broekel and Brenner 2011; Grossman and Helpman 1991), in that we consider a region's strength in knowledge production, i.e. its endowment with tangible knowledge production inputs, to be one of the most crucial factors for a specific position in knowledge networks. In particular, knowledge-intensive organisations (universities, large knowledge-intensive or small highly specialised firms) hold the necessary capacities, not only to engage in collaborations, but also to benefit from knowledge transmitted via network inter-linkages. Since, such organisations tend to be spatially concentrated in urban areas exposing them to a variety of experiences from other organisations, we further assume that regions with higher level of development and knowledge productivity, especially urban regions, are more centrally embedded in FP networks. We call this the *capacity dimension* in determining a region's network embeddedness.

The relational dimension

Studies dealing with the geography of R&D collaboration have shown that there are important non-spatial demarcation mechanisms that influence the formation of R&D relations (Autant-Bernard et al. 2007, Boschma 2005). In a regional context, Scherngell and Barber (2011) highlight that for project based R&D collaborations alternative forms of proximity, such as economic or technological, are even more crucial for establishing and maintaining cross-regional knowledge alliances. For a region's centrality in knowledge networks, this implies that such relational factors might influence network embeddedness in several ways:

Division of labour and increasing need for specialisation in research and innovation, together with tendencies of regional technological clustering and specialisation, let us assume that specialised regions that supply very distinctive technologies and capabilities are in a more favourable position to establish a certain network reputation. They may gain strategic advantages, especially if they are able to distribute those distinct pieces of knowledge required by many others, or attach to central network players due to their unique characteristics. Thus, strongly pronounced high-tech orientation as well as technological specialisation may positively influence a region's strategic position in knowledge networks.

However, innovation literature emphasises that a higher diversification in collaboration partners exposes own research activities to new perspectives, and to different research and application methods (Goetze 2010, Cowan and Jonard 2004). This would imply that regions with broader knowledge production structures more likely take advantage of knowledge pathways that range throughout the entire R&D network, being able to funnel external knowledge that pass through them. Thus, we argue that structural diversification may help to gain control over these divergent knowledge flows, and thus, to better exploit the potential involved in accessing direct and indirect network linkages. We therefore assume that economic and technological diversification places regions in a more central network position. This is referred to as a region's range in knowledge production activities, or the *relational dimension* in determining a region's network embeddedness.

The spatial dimension

The literature on regional knowledge production processes (see, for example, Fischer et al. 2006 and 2009, Fischer und Varga 2003) provides evidence for the influence of properties in neighbouring regions on a region's own knowledge production activity. Thus, we account for external characteristics in surrounding regions, assuming that spatial connectivity affects a region's network position due to spillover mechanisms resulting from economic dependencies, agglomeration dynamics or core-periphery structures in nearby regions (see Feldmann and Kogler 2010, Breschi and Lissoni 2009). This is called the *spatial dimension* in determining a region's network embeddedness.

3 Measuring regional embeddedness in knowledge networks

This section formally specifies the notion of regional network embeddedness from a Social Network Analysis (SNA) perspective. In SNA, it is basically assumed that network interactions act as channels to transmit information and knowledge between actors (Knoke and Young 2008). However, scale and scope of this transmission process may differ considerably with regard to the type of knowledge concerned (Borgatti 2005). Analysing the global structure and dynamics of linkages enables a better understanding of the role that specific actors play in knowledge exchange throughout the network. SNA provides us with a rich analytical toolset to observe structural properties of the network in order to characterise the importance and positioning of specific actors. Actors' centrality is a widely applied concept in this context (Borgatti and Everett 2006, Wassermann and Faust 1994).

In this study, our focus is on different types of knowledge networks across Europe that are the project based R&D collaboration network within the European Framework Programmes (FPs), the co-patent network and the co-publication network. We broadly define network embeddedness in terms of a region's centrality in European knowledge networks. Centrality in each individual network is captured by participations in R&D projects funded by the EU FPs, co-patents and co-publications, respectively. To obtain the regional focus in our analysis, we, first, aggregate individual collaborative activities in time period t to the corresponding regional level. For this purpose, we use a set of i, j = 1,., n=241 NUTS-2 regions (NUTS revision 2003). Second, we construct n-by-n collaboration matrices of the type employed by Scherngell and Barber (2009 and 2011) for FP networks that contain the number of R&D collaborations[1] between two regions i and j, with i = 1,... , n regions in the rows and j = 1,... , n regions in the columns. Note that we construct and analyse the individual networks under consideration separately but introduce our methodological approach in general terms for purposes of readability. The empirical measurement of collaborations in the FPs, co-patenting and co-publications is given in Section 6 of this study.

It is worth noting in this context that organisations rather than regions are the essential actors in knowledge network, and thus, would constitute a more appropriate unit of observation for such kind of analyses. Instead of measuring network centrality at the regional level, an alternative approach would be to define network centrality by means of a bipartite graph directly at the organisational level in a first step (see Barber et al. 2011), and aggregate the observed organisation centralities to the regional level in a second step. However, since we do not have information on the network structures at the organisational for some knowledge

[1] We use full counting procedures for the construction of our collaboration matrices, assigning links for each participating organisation that is located in a different region.

networks under consideration, we stick directly to the regional level. However, we have carried out statistical tests for the FP network – for which we have network structures also at the organisational level – to examine whether inter-regional collaboration matrices serve as appropriate proxies for the underlying network structure at the organisational level. Indeed, results show a high correlation between a region's network centrality calculated at the regional and a region's network centrality calculated at the organisational level and subsequently aggregated to regions[2]. This is a valuable finding, not only in the context of this study, confirming that inter-regional collaboration can serve as appropriate proxies for the underlying network structure at the organisational level, at least when calculating the centrality of a region in a specific network

In terms of graph theory, the n-by-n collaboration matrix for a given knowledge network in year t may be considered as a symmetric n-by-n adjacency matrix of the type

$$\mathbf{A}_t(i,j) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nnt} \end{pmatrix} \tag{1}$$

such that the element a_{ij} contains the collaboration intensity between organisations located in region i and j,) constituting a weighted graph. The unweighted version of the adjacency matrix is given by

$$\mathbf{A}_t^{(bin)}(i,j) = \begin{cases} 0, & a_{ij} = 0 \\ 1, & a_{ij} > 0. \end{cases} \tag{2}$$

that is to be used for measuring specific types of centrality. Further we denote the number of edges incident on a vertex i=1, …, n as the degree k_{it} in a given year t. A path is the alternating sequence of vertices and links, beginning and ending with a vertex, so that the shortest path or geodesic distance g_{ijt} between two regions i and j in time period t is defined as the number of vertices to be passed in the shortest possible path from one vertex to another (see Wassermann and Faust 1994 for further details).

---

[2] For example, Spearmans rank correlation coefficient of eigenvector centrality in the FP network shows a statistically significant value of r_s=0.938 (p < 0.01), for betweenness centrality it is r_s=0.898 (p < 0.01).

Network embeddedness of region i as used in our estimation approach (see Section 6) is captured by two distinct centrality measures, namely betweenness and eigenvector centrality[3]. Each centrality concept is based on different assumptions on how knowledge flows through the network, leading to different perspectives on the positioning of central actors in the networks (Borgatti 2005). The betweenness concept intends to capture the centrality of a node (in our case region) in terms of its position for controlling the flow of information within the network by focusing on the number of shortest paths through this node (region) (Freeman 1979). Thus, central regions benefit from gaining access to various knowledge sources, and, at the same time, take up – independent of their degree – a significant position in influencing the transfer of knowledge within the whole network. In other words, they act as 'gatekeepers' by exerting control over the knowledge flowing through them. In contrast, according to eigenvector centrality a region's centrality depends both on the number and the quality of its connections, assuming that prominent actors act as 'hub' for knowledge transmission and diffusion throughout the entire network. Calculation is based on centralities of all regions in the network in the form of assigning centrality weights that correspond to the average degree of all linked regions (see Bonacich 1987).

We further explore these ideas by providing the mathematical specification of these concepts. For betweenness centrality we utilize the unweighted adjaceny matrix $\boldsymbol{A}_t^{(bin)}$ for a given year t[4]. Thus, in our case, betweenness centrality $y_{it}^{(b)}$ measures how often a region is situated between other, not directly interlinked, regions, in time period t, as defined by

$$y_{it}^{(b)} = \sum_{\substack{j=1 \\ j<q}}^{n} g_{jqt}(i) / g_{jqt} \tag{3}$$

where $g_{jqt(i)}$ is the shortest path between region j and q going through region i at time t, for $i \neq j \neq q$.

---

[3] Further point centrality measures commonly used in SNA are degree and closeness centrality. Degree centrality only focuses on direct links of a vertex, measuring local centrality in particular (see, for example, Wasserman and Faust 1994). In contrast, closeness centrality is based on the shortest distance to all other vertices in the network, indicating how close a distinct vertex is to all other vertices in the network.

[4] We refrain using the weighted version of betweenness centrality, such as for instance defined by Newman (2001), since interpretation of shortest paths in terms of the weighted graphs that we use in this study, that is collaboration intensities between regions, is problematic.

Eigenvector centrality lays – as mentioned above – emphasis on the importance of direct linkages of a vertex in the network, but additionally takes the degree of all other connected vertices into account. Eigenvector centrality $y_{it}^{(ei)}$ of region i at time t is defined to be proportional to the sum of degrees of regions j to which it is connected, using the weighted adjacency matrix $\mathbf{A}_t$:

$$y_{it}^{(ei)} = \frac{1}{\lambda}\sum_{j=1}^{n} a_{ijt} k_{jt} \tag{4}$$

where λ is the largest eigenvalue of $\mathbf{A}_t$[5].

[5] A common notation used in this context is the eigenvector equation as given by $\lambda\ \mathbf{x} = \mathbf{A}\ \mathbf{x}$, where $\mathbf{x}$ is a vector of centralities $\mathbf{x} = (x_1, x_2, ...)$ denoting the eigenvector of the adjacency matrix $\mathbf{A}$ with eigenvalue λ (see Bonacich 1987).

4 Three types of knowledge networks

Our empirical analysis focuses on determinants of a region's network embeddedness in distinct types of European knowledge networks, namely the project based R&D network constituted under the FPs, the co-patent network and the co-publication network. A comparative analysis of the regional determinants of network embeddedness allows us to explicitly take account of specific characteristics involved in the process of knowledge creation as well as the distinctive spatial structure of knowledge diffusion in each network type across Europe.

The FP network is a policy-induced network funded by the EU that spreads over Europe, mainly representing R&D activities at the pre-competitive level (see, for instance, Breschi and Malerba 2009, CEC 2007)[6]. Given the combination of basic and application-oriented research aspects, most FP projects involve research communities of considerable size, directly interlinking heterogeneous, functionally diverse actors across organisational boundaries and different institutional backgrounds (Scherngell and Barber 2011). These self-organised consortia are made up of individual researchers often tied to particular organisations, such as industrial and commercial firms, universities or research organisations, performing collaborative efforts with well-defined objectives in the form of joint R&D projects that are publicly funded on a multi-year basis.

The co-patent network involves industrial research and technological development with a clear application and market orientation (Goetze 2010, Maggioni and Uberti 2009). The network consists of several, often small-scaled and fragmented, inventor communities that have been established for the purpose of sharing knowledge in the (technical) design and development of new technologies. A co-patent is defined as a patent developed by at least two inventors, who are often associated with organisations, representing the visible results of

[6] Since their launch in 1984, the overall objectives of the FPs have been to strengthen the scientific and technological bases of the European scientific community and the European economy to foster international competitiveness, and the promotion of research activities in support of other EU policies (see, for instance, Scherngell and Barber 2009). Funding is open to all legal entities established in the Member States of the European Union – e.g. individuals, industrial and commercial firms, universities, research organisations, etc. – and can be applied by at least two independent legal entities established in different EU Member States or in an EU Member State and an Associated State. Proposals to be funded are selected on the basis of criteria including scientific excellence, added value for the European Community, the potential contribution to furthering the economic and social objectives of the Community, the innovative nature, the prospects for disseminating and exploiting the results, and effective transnational cooperation.

inventive R&D collaboration. In contrast to the policy-induced and project-based FP network, the properties of research partnerships regarding research purpose, time horizon and geographical range of joint inventive endeavours are subject to mechanisms of self-selection. In this sense, co-location of actors is a decisive factor for the first selection of partners. The exchange of tacit and sensitive knowledge components requires high amounts of face-to-face contact and trust, and this may lead in further consequence to path-dependent, persistent and repetitive collaboration structures (Powell and Giannella 2010, Singh 2005).

Co-publications refer to scientific collaborations that mainly involve basic research activities in the academic sphere (Katz and Martin 1997). Co-publications are defined as the product of joint scientific work in the form of co-authored papers in scientific journals. Each listed co-author typically has made substantial contributions to the paper, pointing to some kind of closer interaction during the research process. Due to the increasing specialisation in scientific fields, the need for pooling resources or the increasing importance of interdisciplinary research, researchers are compelled to tie together in research communities consisting of individuals – collaborating scholars and scientists – that are mainly associated with universities or research organisations from different geographical locations (Stephan 2010). In contrast to joint R&D and inventor activities, collaboration mainly takes place within the institutional boundaries of the academic sphere, easing the need for geographical proximity in collaborative efforts (Ponds et al. 2007, Mairesse and Turner 2005). In this sense, coherent scientific standards and codification schemes facilitate first partner selection, especially since mutual research interests and complementaries in knowledge bases could, in principle, be discovered and monitored from the literature. However, the scientific system is to a high degree based on reputation, spurring forces of community demarcation and hierarchical attachment of elite researchers in co-autorship efforts. This leads also to substantial concentration of elite researchers in geographical space (see, for example, Hoekman et al. 2009).

5 Data and explorative analysis

To observe the structure of these three network types, we use data from three different sources. For the FP network we draw on data from the EUPRO database which provides comprehensive information on project-based R&D collaborations funded within the EU FPs[7]. One network link represents a joint R&D project between two organisations. Considering only inter-regional links gives rise to our European network of R&D collaboration at the regional level. In addition, we use the Regpat database to observe the co-patent network containing information on patent applications that have been issued at the European Patent Office (EPO). We use information on the inventor address of an EPO patent application to trace the origin of the invention. This allows us to construct our inter-regional co-patent network, assigning links if a patent application contains at least two inventors located in different regions. Information on co-publications has been obtained from the Web of Science (WoS) database, a bibliographical database indexing journal article worldwide produced by Thomson Reuters. It is one of the most comprehensive sources of information on basic research activities. The WoS contains the institutional addresses of the authors for most articles. In order to reflect inter-regional scientific collaboration, a link in the co-publication network is given when a publication contains two or more authors located in different regions.

Table 1 shows some general characteristics for the three types of knowledge networks. We can observe that the mean degree, i.e. the mean number of collaborations per region, increases in all three network types over the period from 1999 to 2006, but collaboration intensity increases by far more in the FP and co-publication networks than in the co-patent network. The number of nodes and (weighted) edges between them determines structural properties of a network. Thus, constant increase in both sum of links and number of edges within our set of 241 regions or nodes reflects a process of increasing interactions between different regional pairs across Europe for all three networks. Particularly interesting is the comparison of the FP network and the co-publication network. Total link number is higher in the co-publication network, and so are average edge weights. The lower number of edges point to higher

[7] EUPRO has been constructed and maintained by AIT Austrian Institute of Technology. It contains systematic information on project objectives and achievements, project costs, project funding and contract type as well as on the participating organisations including the full name, type of the organisation and geographical location for FP1 to FP7 (see, for instance, Scherngell and Barber 2011).

numbers of collaborations between the same partners in the co-publication than in the FP networks. Furthermore, high values of the standard deviation and degree point to superiority of large research facilities, i.e. research organisations and universities, in the co-publication network. Further, it is remarkable that the inter-regional co-patent network is not only considerably smaller than the FP and the co-publication network, but also much more fragmented[8]. This let us assume that inter-regional co-patenting narrows down to only a few number of strongly tied core nodes in the centre of the network, and numerous weakly connected in the periphery.

Table 1: Network characteristics for the FP, co-patent and co-publication network (1999-2006)

FP	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**
Sum of links	271,772	384,944	474,638	572,764	588,984	804,458	840,452	900,446
No. of edges	14,148	15,837	17,163	17,925	18,221	19,229	19,409	19,602
Mean degree	1,127.69	1,597.28	1,969.45	2,376.61	2,443.92	3,338.00	3,487.35	3,736.29
SD	1,735.51	2,522.76	3,094.08	3,751.08	3,836.74	5,469.81	5,805.03	6,328.92
Min	0	0	0	1	0	0	0	0
Max	16,478	24,858	29,747	36,061	37,076	55,721	60,182	65,597
Graph density	0.49	0.55	0.59	0.62	0.63	0.66	0.67	0.68
Co-patent								
Sum of links	98,404	110,574	112,050	118,168	118,322	127,196	124,630	132,058
No. of edges	3,337	3,656	3,639	3,763	3,770	3,778	3,882	4,018
Mean degree	408.32	458.81	464.94	490.32	490.96	527.78	517.14	547.96
SD	799.12	924.25	950.24	1,016.79	1,022.65	1,115.11	1,054.74	1,077.14
Min	0	0	0	0	0	0	0	0
Max	5,270	6,169	5,984	6,621	7,695	8,960	7,618	7,227
Graph density	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.14
Co-publications								
Sum of links	714,482	747,698	710,034	751,300	854,162	1,089,064	1,179,342	1,392,230
No. of edges	12,611	13,273	13,904	13,956	14,784	15,604	15,708	16,384
Mean degree	2,964.66	3,102.48	2,946.20	3,117.43	3,544.24	4,518.94	4,893.54	5,776.89
SD	5,094.17	5,239.50	4,765.53	4,968.14	5,624.79	7,589.15	7,852.24	9,376.16
Min	0	2	0	3	4	0	2	2
Max	43,364	44,410	38,491	38,588	45,769	64,405	59,030	72,179
Graph density	0.44	0.46	0.48	0.48	0.51	0.54	0.54	0.57

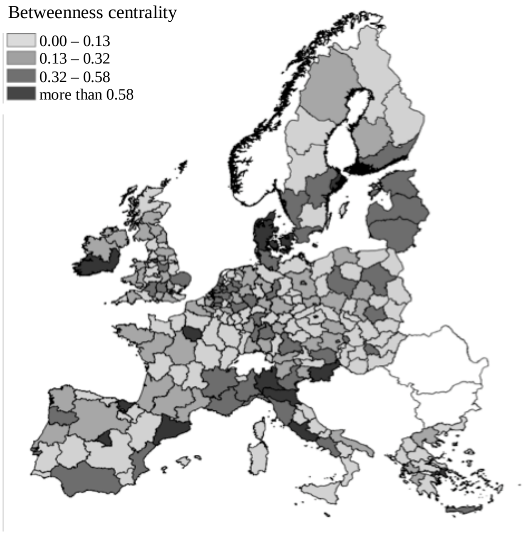
We calculate betweenness and eigenvector centrality for each region *i* as defined in Section 3. Figure 1 visualizes the spatial distribution of the three types of centrality in the year 2006. We can observe considerable differences, on the one hand, between the three types of knowledge networks, and, on the other hand, between our two centrality measures. While spatial

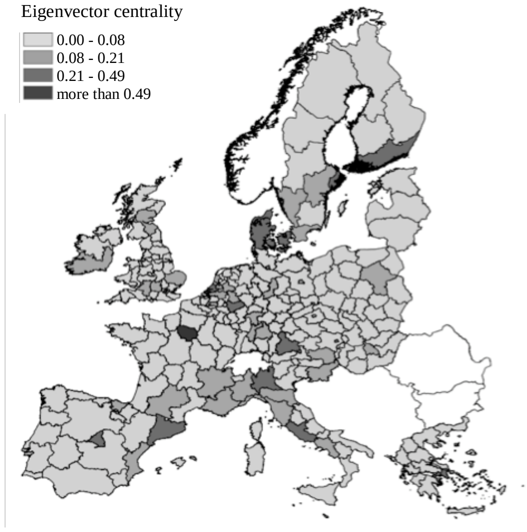
[8] Note that network fragmentation is much higher for the co-patenting network at the organizational level, leading to the fact that centrality measures calculated at the organizational level, especially measures of global network centrality as used in this study, would be subject to considerable bias.

clustering of centralities across neighbouring regions is less significant for the FP network, such spatial clustering becomes apparent in the co-publication network. This spatial concentration becomes even more striking for the co-patent network. Moreover, spatial concentration of network centralities is more pronounced for eigenvector centrality than for betweenness centrality, especially for the co-patent network. Co-inventorship is strongly localized and confined to national – even regional – borders, and, in terms of eigenvector centrality, spatially centred on a single core acting as strong hub for industrial knowledge.

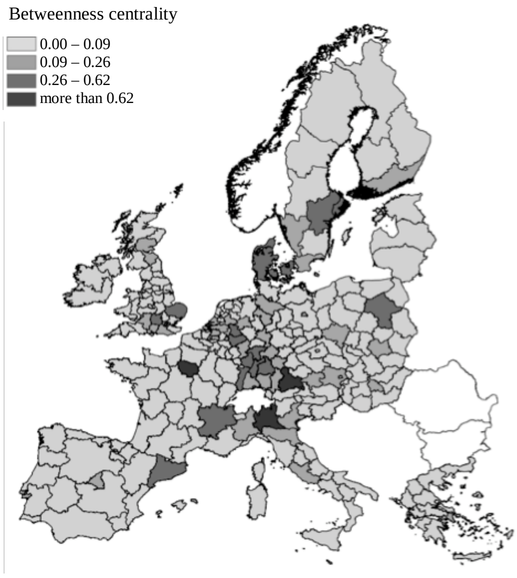
Figure 1: Betweenness and eigenvector centrality for three knowledge networks (2006)

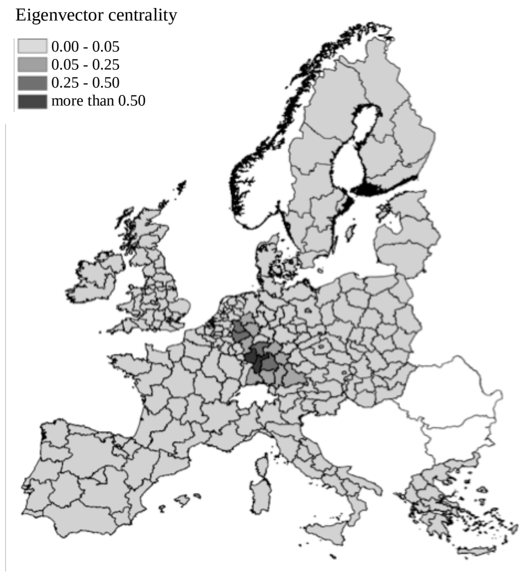
FP Network



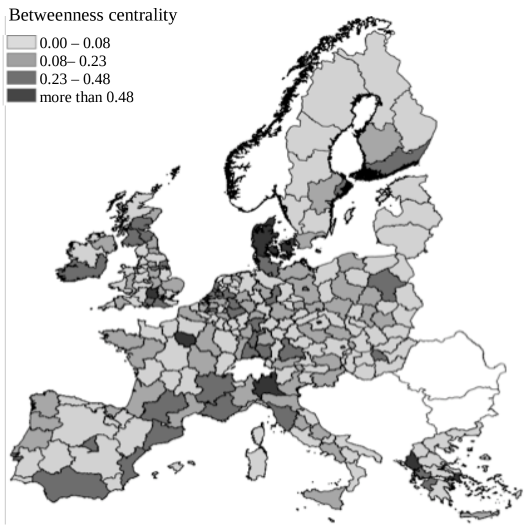


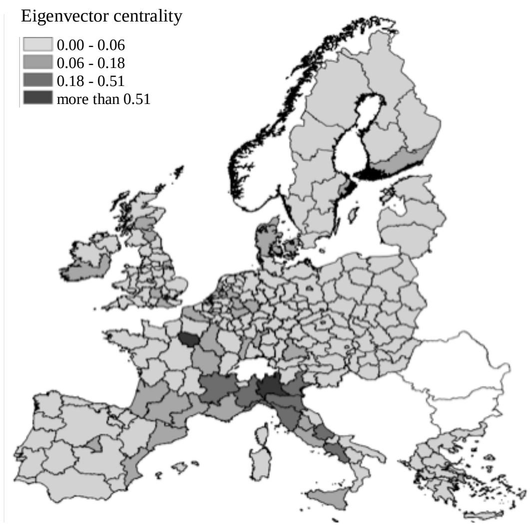
Co-patent network





Co-publication network





Note: Centrality values have been normalised between 0 and 1; natural breaks have been used to classifying data into four categories.

Table 2 provides some spatial autocorrelation analysis using Moran´s I for global spatial autocorrelation to confirm this observation, accompanied by some descriptive statistics. The Moran's I statistically underscores the findings from the spatial visualisation, confirming rather weak and diminishing spatial autocorrelation in our centrality measures for the FP network as compared to the co-publication and especially the co-patent network. For the latter two, we can observe that eigenvector and betweenness centralities are significantly correlated with the centralities in nearby regions. This finding is to be seen in the context of previous studies dealing with spatial autocorrelation of knowledge flows (see, for example, Scherngell and Lata 2012). It is further noteworthy that spatial dependence considerably decreases in the FP network. Table 2 additionally shows that the distribution is highly skewed for both eigenvector and betweenness centrality, pointing to the fact that there are a few – spatially uncorrelated - regions with relatively high centrality values, while most regions show comparatively low centralities.

Table 2: Descriptive Statistics of regional centrality by knowledge network

Betweenness	**FP 1999**	**FP 2006**	**Co-patents 1999**	**Co-patents 2006**	**Co-pub 1999**	**Co-pub 2006**
Morans'I	0.075**	-0.006	0.370***	0.316***	-0.046	-0.047
Mean	57.56	53.63	123.47	112.83	68.86	67.73
Median	27.71	29.47	29.26	26.26	34.71	35.20
Stand. Dev.	73.86	67.87	230.15	210.99	88.53	83.35
Skewness	2.10	2.24	3.59	3.54	2.07	2.00
Kurtosis	5.22	6.07	17.26	16.51	5.19	4.62
Eigenvector						
Morans'I	0.301***	0.025	0.544***	0.524***	0.182***	0.210***
Mean	0.07	0.07	0.03	0.03	0.04	0.05
Median	0.03	0.03	0.00	0.00	0.01	0.01
Stand. Dev.	0.11	0.11	0.11	0.13	0.10	0.10
Skewness	3.69	3.89	6.41	5.89	5.52	5.55
Kurtosis	21.40	23.83	45.20	36.11	40.36	40.26

6 Modelling a region's centrality in R&D networks

The exploratory and descriptive analysis has revealed interesting spatial patterns of regional network centralities across Europe. However, the question that arises is which regional characteristics drive the observed spatial patterns. At this point of the study we are interested in measuring the regional determinants of embeddedness in European knowledge networks. In order to estimate the relationship between network embeddedness and region-internal and region-external characteristics we implement a panel version of the Spatial Durbin Model (SDM)[9] in the form of

$$y_t = \delta W y_t + X_t \beta + W X_t \gamma + u_t \qquad t = 1,\ldots, T=8 \qquad (5)$$

with

$$X_t = [C_t \;\; Z_t] \qquad (6)$$

where y_t=(y_{1t}, …, y_{nT}), X_t=(X_{1t}, …., X_{nT}) and u_t =(u_{1t}, …, u_{nT}). y_t is a n-by-1 vector consisting of observations y_{it} (i, …., n =241; t=1, ..., T=8) corresponding to the centrality of region i in a knowledge network at time t. Note that we have specified network centrality in terms of betweenness centrality $y_{it}^{(b)}$ (Eq. 3) or eigenvector centrality $y_{it}^{(ei)}$ (Eq. 4). X_t is an n-by-K matrix of observations on the explanatory variables, including sets of C_t (n-by-S; s=1, ..., S=4) variables indicating the capacity dimension of knowledge production within a region, and Z_t (n-by-Q; q=1, ..., Q=3) indicating the relational dimension of knowledge production. β is a K-by-1 vector of the regression parameters associated with the K explanatory variables.

The non-stochastic, time-invariant n-by-n spatial weights matrix W gives the spatial configuration of our set of regions with w_{ij} being the (i,j)th element of W, such that it contains elements of the form w_{ij} = 1 if region i and region j are spatial neighbours and share a common border, and w_{ij} = 0 otherwise[10]. Wy_t is a n-by-1 vector representing the average of the spatially lagged network centrality values in neighbouring regions of i, δ is the spatial autoregressive coefficient measuring the strength of the spatial autoregressive relation between neighbouring regions. The n-by-K matrix WX_t includes the spatially weighted observations on the K explanatory variables of neighbouring regions, with γ being a K-by-1 vector of parameters capturing spatial interaction effects induced by region-external characteristics. u_t is the disturbance term for time period t.

---

9 The spatial dependence relation in our observations on network centrality as indicated by the Moran' I statistic provide strong motivation for the SDM. Note further that the SDM nests conventional spatial regression model specifications, such as spatial lag (SAR) or spatial error (SEM) models (see LeSage and Pace 2009 for details). In order to unravel the true data-generating process and decide on - besides the conceptual motivation – appropriate model specification, a general-to-specific approach has been applied confirming our SDM specification against more specific SAR or SEM specifications.

10 We use a row standardized version of W allowing interpretation of the spatial lags of the independent variables being the weighted average impact on region i by their neighbouring regions. Note further that the meaning of neighbourhood is used in the sense of spatial relatedness throughout this study.

Given the panel nature of our model, two-way error component disturbances may be used to jointly account for unobservable space- and time-specific effects (Baltagi 2008). Thus, we specify the disturbance term u_t in the form of

$$u_t = \mu + \tau_t + \varepsilon_t \tag{7}$$

where μ denotes unobservable random time-invariant variability between regions, and τ_t is the counterpart for omitted time-specific region-invariant fixed effects. ε_t denotes a n-by-1 vector of the residual error term (*iid*) for time period t with zero mean and variance σ^2. Eq. (7) is labelled the mixed effects panel SDM specification, including time-specific fixed effects and region-specific random effects. We use within transformation to obtain time fixed effects estimates for our observations in y_t and X_t that allows us to account for time-dependent but region-invariant variability in our model specification (see Park 2011, Baltagi 2008, Elhorst 2003). Accounting for time-specific shocks is particularly important when analysing policy-induced network structures such as given by the EU FPs. The region-specific effect μ is specified as a random component with $\mu \sim N(0,\sigma_\mu^2)$ and independent of $u_{it.}$ The random effects approach seems to be appropriate given our extended set of explanatory variables that capture regional knowledge production in terms of its direction, structure and strength. In fixed effects models, region-specific effects tend to absorb significance of structure-indicating, relatively time-variant variables[11]. Moreover, in measuring knowledge network centrality at the regional level, unobservable effects are likely to be related with individual, organisation-specific decisions to engage in collaborations (see, for example, Paier and Scherngell 2011, Autant-Bernard et al. 2007), which we assume to follow a random pattern across European regions. We use Maximum Likelihood estimation procedures to estimate the parameters in the SDM specification (see Elhorst 2003 for details).

The advantage of the SDM in terms of model specification is that - due to including spatial lags of both the dependent and explanatory variables - the risk of inconsistent and biased estimates caused by correlated but omitted explanatory variables is limited as compared to restricted SAR or SEM specifications (Le Sage and Pace 2009). The most noteworthy benefits of using SDM specification is that it allows disclosing the influence of various channels of

[11] Note further that the spatial fixed effects model could not be estimated consistently due to relatively time-invariant independent variables. Further, the number of observations for n=241 regions, in contrast to t=8 time periods, is relatively large (Elhorst 2003, Baltagi 2008).

spatial spillovers effects from neighbouring regions, and thus, demarcating effects induced by each explanatory variable (Autant-Bernard 2012). However, interpretation and statistical inferences for direct and indirect effects are not straightforward since spatial dependence of regional observation may cause feedback loops that walk through the dependent and explanatory variables of neighbouring regions. Le Sage and Pace (2009) provide valuable approaches – also used in this study – for assessing the magnitude and statistical inferences for the direct, region-internal, and indirect, region-external spillover effects of the explanatory variables.

Variables that explain a region's embeddedness

In modelling a region's embeddedness in knowledge networks, we consider two different types of explanatory variables X=[C Z]. The variables in C (s=1, ..., S=4) account for aspects of a region's knowledge production capacities (*capacity dimension*) in form of

i. C_1 capturing total regional R&D expenditures (log of public and private R&D expenditures in % of GRP), used as a proxy for the magnitude of financial inputs in knowledge production.

ii. C_2 being the logarithmic share of population with tertiary education (corresponding to levels 5 and 6 of the ISIC 1997 classification system), measuring a region's endowment with human capital. Human capital not only indicates the potential of performing knowledge intensive R&D, but also conditions a region's absorptive capacity that is regarded as one of the major necessities to collaborate and reap full benefits of joint R&D.

iii. C_3 that is the logarithmic form of the gross regional product (GRP) per capita, proxying the economic development, productivity and general socio-economic potential of a region that are assumed to be decisive stimuli for a region's R&D performance.

iv. C_4 denoting the region's population density as measured by the number of inhabitants per square kilometre, used as proxy variable for the degree of urbanisation and for agglomeration effects in urban areas.

Our set of Z variables (q=1, ..., Q=3) includes proxies for a region's range of knowledge production activities (*relational dimension*). Here,

i. Z_1 captures the degree of technological specialisation within region i, using an index of specialisation of a region's patent portfolio[12]. The focus of a region's knowledge production activities may range over a variation of different fields, or be very specific and specialised in a certain technology. As joint R&D means to pool, relate and assemble very different knowledge components, a region's technological orientation may influence network embeddedness. High specialisation may probably facilitate achieving a hub position within a network, while technological diversification may determine the ability to engage in several networks, reflecting the opportunity to exploit and combine inputs from different knowledge sources.

ii. Z_2 denotes regional high-tech orientation measured by the number of high-tech patents per million employees, used in logarithmic form. This variable is a proxy for technological strength of a region's knowledge base. We further assume that high-tech industry facilities contribute to higher absorptive capacity of knowledge transmitted via R&D collaborations.

iii. Z_3 is the degree of industrial diversity within region i measured in terms of an industrial diversity index[13]. We include this variable since we assume that knowledge production activity benefits from a diversified industrial structure. Inter-sectoral production chains and interdependencies between economic sectors affect the range of knowledge production activities, notably in terms of backward and forward linkages, interdisciplinarity and applicability of R&D. Moreover, a diverse economic structure may enhance a region's attractiveness in a range of collaboration formations, extending opportunities to access different knowledge sources.

Data on most independent variables have been drawn from the Eurostat regional database, while information on patents was taken from the European Patent Office (EPO) database.

---

12 The index is defined by $c_{it}^{(4)} = \frac{1}{2}\sum_p \left|s_{ip} - \overline{s}_p\right|$ where s_{ip} is the region's i share of patents in a specific IPC class p and $\overline{s}_p$ is the mean of IPC class p. Patents were taken into account at a three-digit level corresponding to the International Patent Classification (IPC).

13 We include five different main economic sectors, namely agriculture, manufacturing, construction, private services and non-market service sector. The index of specialisation to account for industrial diversity is defined as $z_{it}^{(1)} = \frac{1}{2}\sum_p \left|o_{ip} - \overline{o}_p\right|$ where o_{ip} is the region's i share of gross value added in a specific sector p (indexed p = 1, …., 5) and $\overline{o}_p$ is the mean of sector p for n=241 regions.

7 Estimation results

This section discusses the estimation results for our regional network embeddedness models as specified by Eqs. (5)-(7). Table 3 reports the Maximum Likelihood (ML) parameter estimates and the associated *t*-statistics for the mixed effects panel version of the SDM. The first column contains results for betweenness centrality (BC), and the second column for eigenvector centrality (EC) for each network type.

The bottom of Table 3 provides specification tests as well as various goodness-of-fit measures. Significant robust LM tests and Likelihood Ratio (LR) tests suggest the presence of spatial dependence in our empirical setting, confirming not only implementation of a spatial regression model, but also the SDM specification against spatial lag (SAR) or spatial error (SEM) specifications. Thus, geographical space in terms of regional interaction effects does matter in explaining embeddedness in European knowledge networks. In this context, a first remarkable finding is that estimates for the spatial autoregressive parameter (δ) are highly significant and positive for the FP network and the co-patent network but not for the co-publication network.

Tables 4, table 5 and table 6 present average impact estimates for each network on the magnitude and significance of direct, indirect and total impacts on a region's network position that would arise from a change in one unit of our regional characteristics, averaged over space and time (note that most estimates of our explanatory variables are interpretable as elasticities). More specifically, the direct impacts give the effects of region-internal characteristics on a region's network embeddedness, while the indirect impacts estimates report the sum of spatial spillover effects, i.e. influences of region-external characteristics, in our regional setting[14]. Then, the overall influence of distinct characteristics on regional network embeddedness at the regional level is given by the total impact estimates.

Table 4 presents scalar summary impact measures for the FP network. Concerning direct impacts, we can observe that a region's *capacity* in knowledge production, in particular R&D expenditures and economic strength, is decisive for a region's embeddedness in the FP network, in terms of both betweenness and eigenvector centrality. Direct impact estimates

[14] Differences between SDM coefficient estimates and impact estimates give indication on the magnitude of the feedback effects.

further suggest that well-educated human capital is a significantly important capacity factor for achieving betweenness centrality, and thus, for performing a gatekeeper function in the network (note that for these impacts the confidence interval at the 95 percent level do not span zero).

Table 3: ML estimation results for the mixed effects panel Spatial Durbin Model (SDM)

	FP		Co-Patent		Co-Publication	
	BC	**EC**	**BC**	**EC**	**BC**	**EC**
RD expenditures [$\beta_1^{(c)}$]	0.443** (-2.354)	0.280*** (3.950)	1.710*** (6.162)	0.838*** (6.570)	0.395*** (2.754)	0.108** (2.305)
Human capital [$\beta_2^{(c)}$]	2.160*** (4.321)	0.042 (0.220)	3.412*** (4.785)	1.627*** (4.883)	0.733* (1.915)	0.242* (1.870)
GRP p.c. [$\beta_3^{(c)}$]	1.006** (2.400)	1.291*** (7.492)	0.826 (1.475)	0.308 (1.143)	-0.045 (-0.141)	0.082 (0.652)
Population density [$\beta_4^{(c)}$]	0.001 (1.393)	0.001*** (3.683)	0.000 (0.960)	0.000 (0.915)	0.001*** (2.672)	0.001*** (3.480)
Technol. specialisation [$\beta_1^{(z)}$]	0.254*** (4.399)	0.078*** (3.834)	-0.415*** (-4.151)	-0.168*** (-3.956)	0.055 (1.262)	-0.005 (-0.361)
High-tech patents [$\beta_2^{(z)}$]	0.056 *** (4.587)	0.018*** (4.298)	0.214*** (9.556)	0.054*** (5.876)	0.023** (2.517)	0.006** (2.427)
Industrial diversity [$\beta_3^{(z)}$]	0.048** (1.378)	0.036** (2.495)	0.112*** (2.647)	0.066*** (3.106)	0.082*** (3.061)	0.035*** (3.217)
w. RD expenditures. [$\gamma_1^{(c)}$]	-0.928*** (-2.790)	-0.286** (-2.233)	0.697 (1.427)	0.015 (0.067)	-0.374 (-1.468)	-0.113 (-1.294)
w. Human capital [$\gamma_2^{(c)}$]	-1.314* (-1.893)	0.074 (0.273)	-3.959*** (-4.136)	-1.477*** (-3.251)	-0.948* (-1.784)	-0.148 (-0.797)
w. GRP p.c. [$\gamma_3^{(c)}$]	0.725*** (3.152)	0.111 (1.153)	1.075 (3.534)	0.684*** (4.701)	0.595*** (3.361)	0.257*** (3.100)
w. Pop. density [$\gamma_4^{(c)}$]	-0.001 (-1.062)	-0.001*** (-2.921)	0.000 (0.393)	0.000 (-1.243)	-0.001 (-1.044)	0.000 (-0.668)
w. Technol. spec. [$\gamma_1^{(z)}$]	-0.071 (-0.650)	0.022 (0.547)	0.507*** (2.786)	-0.176** (-2.216)	-0.111 (-1.338)	-0.054** (-2.146)
w. High-tech pat. [$\gamma_2^{(z)}$]	0.378 (1.542)	0.027*** (3.053)	0.153*** (3.480)	0.086*** (4.737)	-0.004 (-0.204)	0.002 (0.393)
w. Ind. diversity [$\gamma_3^{(z)}$]	-0.069 (-1.378)	-0.080*** (-3.735)	-0.049 (-0.804)	-0.042 (-1.355)	-0.068* (-1.731)	0.008 (0.464)
SAR coefficient [δ]	0.105*** (3.378)	0.271*** (9.574)	0.065* (2.070)	0.190*** (6.377)	-0.044 (-1.353)	0.029 (0.907)
Model diagnostics and model fit						
Teta	0.206	0.138	0.407	0.295	0.196	0.072
Log Likelihood	-4157.63	-2200.73	-5328.77	-3592.22	-3618.70	-1400.26
LR Test for SAR	50.44***	172.97***	80.06***	80.06***	37.17***	178.29***
SEM	68.63***	163.99***	83.62***	83.62***	47.73***	186.29***
Robust LM test for SAR	58.64***	81.88***	4.940**	28.32***	51.43***	23.29***
SEM	65.89***	127.72***	0.150	135.40***	48.34***	138.95***

Note: *t*-values are given in brackets. Mixed effects refer to time-specific fixed effects and region-specific random effects estimation. BC denotes betweenness centrality, EC denotes eigenvector centrality. The independent variables are defined as given in the text. LR Test is the Likelihood Ratio Test for the SAR and SEM; robust LM Test is the Lagrange Multiplier Test for the SAR and SEM. We checked variance inflation factors (VIFs) for the variables range inferring that there are no multicollinearity problems. ***significant at the 0.01 significance level, **significant at the 0.05 significance level, *significant at the 0.1 significance level.

Concerning our set of variables reflecting *relational* aspects of regional knowledge production activities, all direct impact estimates are significant and slightly positive for eigenvector centrality. Industrial diversity, technological specialisation and, in particular, high-tech orientation seem to reinforce each other in promoting knowledge hubs in the FP network, Knowledge production capacities prove to be the primary condition though. The same holds true for betweenness centrality, even if industrial diversity has no direct impact.

Table 4: Average impacts estimates for the FP network

	Betweenness Centrality			Eigenvector Centrality		
Direct impacts	**Lower**	**Mean**	**Upper**	**Lower**	**Mean**	**Upper**
RD expenditures	**0.052**	**0.424**	**0.797**	**0.125**	**0.267**	**0.408**
Human capital	**1.172**	**2.142**	**3.112**	-0.327	0.047	0.422
GRP p.c.	**0.190**	**1.019**	**1.847**	**0.982**	**1.316**	**1.650**
population density	0.000	0.000	0.001	0.000	0.000	0.001
Technol. Spec.	**0.141**	**0.251**	**0.361**	**0.040**	**0.080**	**0.120**
High-tech patents	**0.032**	**0.056**	**0.081**	**0.012**	**0.021**	**0.029**
Industrial diversity	-0.020	0.046	0.112	**0.004**	**0.032**	**0.059**
Indirect impacts	**Lower**	**Mean**	**Upper**	**Lower**	**Mean**	**Upper**
RD expenditures	**-1.618**	**-0.930**	**-0.242**	-0.588	-0.266	0.056
Human capital	-2.513	-1.138	0.238	-0.512	0.107	0.727
GRP p.c.	**0.396**	**0.863**	**1.330**	**0.318**	**0.577**	**0.837**
population density	-0.002	-0.001	0.001	-0.002	-0.001	0.000
Technol. Spec.	-0.275	-0.047	0.182	-0.047	0.054	0.155
High-tech patents	**-0.004**	**0.045**	**0.095**	**0.018**	**0.040**	**0.061**
Industrial diversity	-0.167	-0.069	0.028	**-0.138**	**-0.089**	**-0.039**
Total impacts	**Lower**	**Mean**	**Upper**	**Lower**	**Mean**	**Upper**
RD expenditures	-1.298	-0.506	0.286	-0.364	0.001	0.366
Human capital	-0.237	1.004	2.245	-0.474	0.155	0.783
GRP p.c.	**1.023**	**1.881**	**2.739**	**1.455**	**1.894**	**2.332**
population density	-0.001	0.000	0.001	-0.001	-0.001	0.000
Technol. Spec.	-0.054	0.204	0.463	**0.018**	**0.134**	**0.251**
High-tech patents	**0.047**	**0.102**	**0.157**	**0.035**	**0.060**	**0.085**
Industrial diversity	-0.120	-0.023	0.073	**-0.107**	**-0.057**	**-0.007**

Note: Significant estimates using 95 per cent confidence intervals are marked in bold.

Moreover, positive *spatial* spillover impacts on a region's network embeddedness can only be observed for the estimates for GRP per capita and for high-tech orientation of neighbouring regions. Thus, a region's FP network embeddedness seems to benefit from inter-regional knowledge spillovers in high-tech industries. In contrast, spillover effects from R&D expenditures are significantly negative, and even of larger magnitude than their direct effect counterparts, pointing to negative dependencies of knowledge creation activities in geographic space that might be due to pulling and polarisation forces caused by the presence of important research facilities or knowledge-intensive firms.

Given the total impact estimates for our regional characteristics in the FP network model version, the pattern observed for direct effects partly has reversed when taking a pan-European perspective. Apart from a region's GRP per capita, only relational knowledge creation factors in terms of high-tech orientation for betweenness centrality, as well as technological and industrial specialisation for eigenvector centrality seem to significantly influence network positioning in the FP network.

Table 5: Average impact estimates for the co-patent network

	Betweenness Centrality			Eigenvector Centrality		
Direct impacts	**Lower**	**Mean**	**Upper**	**Lower**	**Mean**	**Upper**
RD expenditures	**1.191**	**1.718**	**2.245**	**0.593**	**0.839**	**1.085**
Human capital	**1.924**	**3.340**	**4.755**	**0.963**	**1.587**	**2.211**
GRP p.c.	-0.257	0.851	1.959	-0.204	0.337	0.878
Population density	0.000	0.000	0.001	0.000	0.000	0.000
Technol. Spec.	**-0.604**	**-0.407**	**-0.209**	**-0.258**	**-0.176**	**-0.093**
High-tech patents	**0.173**	**0.217**	**0.260**	**0.041**	**0.058**	**0.076**
Industrial diversity	**0.033**	**0.111**	**0.189**	**0.026**	**0.065**	**0.104**
Indirect impacts	**Lower**	**Mean**	**Upper**	**Lower**	**Mean**	**Upper**
RD expenditures	-0.138	0.820	1.778	-0.284	0.201	0.686
Human capital	**-5.748**	**-3.763**	**-1.779**	**-2.257**	**-1.337**	**-0.418**
GRP p.c.	**1.750**	**1.140**	**0.530**	**0.550**	**0.852**	**1.154**
population density	-0.001	0.000	0.002	-0.001	-0.001	0.000
Technol. Spec.	**0.119**	**0.480**	**0.842**	**-0.410**	**-0.241**	**-0.071**
High-tech patents	**0.083**	**0.169**	**0.254**	**0.070**	**0.111**	**0.152**
Industrial diversity	-0.164	-0.045	0.074	-0.098	-0.033	0.031
Total impacts	**Lower**	**Mean**	**Upper**	**Lower**	**Mean**	**Upper**
RD expenditures	**1.451**	**2.538**	**3.625**	**0.471**	**1.041**	**1.610**
Human capital	-1.976	-0.424	1.129	-0.587	0.250	1.086
GRP p.c.	**0.903**	**1.991**	**3.079**	**0.578**	**1.190**	**1.190**
population density	-0.001	0.001	0.002	-0.001	0.000	0.000
Technol. Spec.	-0.309	0.074	0.456	**-0.608**	**-0.416**	**-0.224**
High-tech patents	**0.293**	**0.385**	**0.477**	**0.123**	**0.169**	**0.216**
Industrial diversity	-0.046	0.066	0.178	-0.034	0.032	0.097

Note: Significant estimates using 95 per cent confidence intervals are marked in bold.

For the co-patent network (Table 5), the pattern is more homogenous for betweenness and eigenvector centrality, pointing to more pervasive regional determinants of network embeddedness in the inventor network. Here again two *capacity* factors are the most decisive region-internal drivers, namely financial and human resources. A well-educated labour force is even twice as important as internal R&D expenditures for regions to gain high reputation in inter-regional co-inventorship across Europe. Interestingly, the influence of a region's GRP per capita is insignificant. These findings let us assume that centrality in the co-patent network is much more related with the factual strength in knowledge production activities, than with economic and agglomeration strength that, for example, highly determine centrality in the R&D collaboration network.

However, *relational* factors indicating a region's range in knowledge creation activities is equally important for reaching high betweenness or eigenvector centrality in the co-patent network. More specifically, a region's high-tech orientation and industrial diversity foster co-patent network centrality, while technological specialisation will hamper it, and this holds true for both, acting as a hub or knowledge distributor in the network as well as performing gatekeeper or brokering activities for several weakly connected actors. In both positions diversification, and associated with this, the necessary ability to absorb and exploit very different pieces of knowledge seem to be of particular importance.

The most interesting finding regarding the influence of *spatial* spillover effects on network centrality in the co-patent network concerns the negative estimate for region-external human capital, indicating that a high level of human capital in nearby regions decreases a region's ability to gain network centrality. Such kind of spatial interaction effects in research and innovation activities are most likely due to spatial clustering of human capital and the restricted availability of human resources. Thus, concentration of highly educated people in neighbouring regions, for example due to the presence of research infrastructure or knowledge-intensive firms, seems to impede central positioning in co-patent networks.

Table 6: Average impact estimates for the co-publication network

	Betweenness Centrality			Eigenvector Centrality		
Direct impacts	**Mean**	**Lower**	**Upper**	**Mean**	**Lower**	**Upper**
RD expenditures	**0.398**	**0.122**	**0.675**	**0.107**	**0.011**	**0.203**
Human capital	0.735	-0.021	1.491	0.241	-0.013	0.495
GRP p.c.	-0.044	-0.664	0.576	0.081	-0.180	0.342
population density	0.001	0.000	0.001	0.001	0.000	0.001
Technol. Spec.	0.057	-0.030	0.143	-0.005	-0.029	0.020
High-tech patents	**0.023**	**0.005**	**0.041**	**0.006**	**0.001**	**0.011**
Industrial diversity	**0.084**	**0.030**	**0.138**	**0.034**	**0.013**	**0.055**
Indirect impacts	**Mean**	**Lower**	**Upper**	**Mean**	**Lower**	**Upper**
RD expenditures	-0.369	-0.829	0.091	-0.109	-0.276	0.058
Human capital	-0.888	-1.897	0.121	-0.135	-0.489	0.219
GRP p.c.	**0.549**	**0.207**	**0.890**	**0.257**	**0.092**	**0.422**
population density	-0.001	-0.002	0.000	0.000	-0.001	0.000
Technol. Spec.	-0.108	-0.265	0.050	**-0.053**	**-0.101**	**-0.005**
High-tech patents	-0.005	-0.040	0.030	0.002	-0.008	0.013
Industrial diversity	-0.067	-0.142	0.009	0.009	-0.022	0.040
Total impacts	**Mean**	**Lower**	**Upper**	**Mean**	**Lower**	**Upper**
RD expenditures	0.030	-0.490	0.549	-0.002	-0.192	0.189
Human capital	-0.153	-0.960	0.654	0.106	-0.236	0.448
GRP p.c.	0.505	-0.055	1.064	**0.338**	**0.082**	**0.595**
population density	0.000	0.000	0.000	0.000	0.000	0.001
Technol. Spec.	-0.051	-0.217	0.115	**-0.058**	**-0.110**	**-0.005**
High-tech patents	0.018	-0.021	0.056	0.009	-0.003	0.020
Industrial diversity	0.017	-0.048	0.083	**0.043**	**0.013**	**0.073**

Note: Significant estimates using 95 per cent confidence intervals are marked in bold.

For network embeddedness in the co-publication network we observe similar magnitude and significance between our impact estimates (Table 6) and our SDM coefficient estimates (Table 3), confirming the minor relevance of feedback effects in co-publication activity between nearby regions. Furthermore, embeddedness in inter-regional co-publication networks is only marginally determined by the set of *capacity*-specific or *relational* factors. While the significant impact exerted by financial R&D inputs is comparatively high in magnitude, a region's high-tech orientation as well as its industrial diversification has only minor effects on a region's position in the scientific collaboration network. Remarkable in this context is that human capital (measured by the share of labour force with tertiary education) does not affect a region's embeddedness in co-publication networks. However, this observation might be due to relatively low employment intensity in the scientific sector as compared to other knowledge-intensive sectors (such as engineering or specific services).

In addition, we find that *spatial* interaction effects arising from our explanatory variables are of minor significance for co-publication network embeddedness than for centrality in the FP or co-patent network. The sole exception is that economically powerful surrounding regions positively influence betweenness and eigenvector centrality, pointing to the significance of economically based spillover relations for centrality in scientific networks, for example due to cross-regional commuting.

Total impact estimates for the co-publication model version show only significant and positive effects for a region's GRP p.c. as well as its industrial and technological diversification structure. Thus, positive capacity-based effects due to the level of R&D expenditures will be removed when region-internal and counteracting external impacts are considered together in our regional arrangement. For betweenness centrality we do not observe any significant total impacts arising from regional characteristics at all. Thus, regional factors seem to be far more important in the FP and co-patenting network, and affect co-publication only marginally. This finding let us assume that scientific collaboration is more likely to be based on the relations of individual researchers, and thus, to a lesser degree determined by the characteristics of the regional environment where the scientists are embedded. Moreover, the results provide evidence that higher degree of codification in scientific knowledge makes spatial relations to be less significant in scientific collaboration. Furthermore, we find no significant direct or indirect impacts of population density on a region's centrality in the different knowledge networks, providing evidence that urban agglomeration is not a sufficient driving force for knowledge network embeddedness of a region.

8 Discussion and concluding remarks

This study focuses on a region's embeddedness in European knowledge networks, taking a comparative perspective on the inter-regional R&D collaboration network as given by the European Framework Programmes (FPs), the inter-regional co-patent network and the inter-regional co-publication network. We aimed at estimating how regional characteristics affect a region's network centrality in each distinct network type. In order to take account of the distinctive structural properties of each network, we applied a Social Network Analysis (SNA) perspective, defining a region's network embeddedness in terms of its centrality according to inter-regional collaboration intensities. By this, we used eigenvector centrality, placing the region in a central hub position, and betweenness centrality, assigning central regions the ability to control knowledge diffusion in the network. In modelling network centrality we distinguished between characteristics that reflect a region's knowledge production capacities (*capacity dimension*), and a region's range of knowledge production activities (*relational dimension*). Our empirical model specification in form of a mixed effects panel version of the Spatial Durbin Model (SDM) allows accounting for the *spatial dimension* in explaining network embeddedness by differentiating between region-internal and region-external impacts arising from our set of explanatory characteristics.

From a comparative perspective, our empirical analysis of the FP, the co-patent and the co-publication network spanning across European regions produced interesting commonalities between these types of networks, but also striking differences that might be traced back to the very distinctive knowledge generation processes:

First, we found evidence that network structure is highly determined by the spatial structure of inter-regional collaboration in each network type. While we can observe comparable patterns in terms of network size, regional collaboration intensity and the geographic location of the most central regions in the FP network and the co-publication network, structural properties of the co-patent network are strikingly different. Spatial distribution of network centrality, in particular eigenvector centrality, reflects the highly localised and fragmented knowledge creation structures in co-inventorship. Both phenomena, i.e. spatial concentration of collaboration and network fragmentation, more likely restrict knowledge diffusion across Europe in the co-patent network than for the FP project and co-publication network. The less skewed distribution of centrality in the FP network seems to reflect EU policy goals related with the closer integration of R&D across Europe. Spatial distribution of co-publication network centrality seems to be as equally distributed across the European scientific landscape, though centrality is more concentrated on European scientific core regions, even more reflecting clustering tendencies of powerful scientific players with high reputation.

Second, we revealed that regional characteristics in terms of *capacity*-specific (financial and human R&D resources, economic strength and urban agglomerations) and *relational* (technological specialisation, high-tech orientation and industrial diversity) factors deliver important insights for explaining knowledge network embeddedness at the regional level. Although a region's own strength in knowledge production (direct impact) proves to be one of the most essential regional determinants, different aspects are addressed in order to achieve a central position in the distinct network types. While in the politically induced FP network a region's economic potential is significantly important, it is human capital in the co-patent network and financial R&D resources in the co-publication network that accounted for the most significant impact. To explain these differences, the focus of knowledge creation in the distinct network types needs to be considered; the FP network combines basic with applied research inputs arising from several sources, in contrast to the co-patent network where industrial application-oriented, often tacit, elements determine inventive activities, while the co-publication network involves mainly basic research activities in the academic field.

We further showed that *relational* factors are of minor significance for network centrality than *capacity*-indicating factors. High-tech orientation and industrial diversity seem to be almost equally conducive factors for all three types of knowledge networks, while technological specialisation seems not to be a significant criterion for centrality in the co-publication network. In this regard again, distinctive knowledge creation foci seem to find expression.

Third, we provided empirical evidence that the *spatial dimension*, i.e. indirect, spatial spillover impacts from neighbouring regions, significantly influence centrality in knowledge networks. Network embeddedness in the co-patent network is subject to spatial spillovers most likely due to higher localisation of knowledge creation in inventor activities. We found positive inter-regional interdependencies for economic strength and high-tech orientation, while knowledge creation capacity seems to exert negative impacts in neighbouring regions' network embeddedness. Thus, we can assume that positive extra-regional knowledge externalities might prevail over negative regional interaction effects as long as no considerable pulling forces are at work that directly address restriction in (financial or human) resource endowment. Moreover, we found significant spatial spillovers of knowledge creation capacities only for R&D collaboration or co-inventor activities, while centrality in the scientific collaboration network seems not to be affected by spatial dependencies in knowledge production structures of nearby regions.

Some ideas for further research come to mind: *First*, sector- specific analyses as well as further statistical information on regional peculiarities will enrich our understanding of how regional conditions are related to inter-regional R&D network embeddedness. *Second*, from a methodological point of view, improvements regarding the quantification of spillovers (direct and indirect) passing through R&D networks may be valuable to enrich our understanding on the relational and spatial structure of knowledge diffusion. *Third*, it may be of particular interest how embeddedness in inter-regional research collaboration affects not only knowledge production but also economic performance of locations, requiring advanced theoretical and modelling approaches to establish the link between inter-regional knowledge diffusion, localised exploitation and economic impacts.

References

Asheim, B., Lawton Smith, H. and Oughton, C. (2011): Regional Innovation Systems: Theory, Empirics and Policy. *Regional Studies* 45 (7), 875–891

Autant-Bernard, C. (2012): Spatial Econometrics of Innovation: Recent Contributions and Research Perspectives. *Spatial Economic Analysis* 7 (4), 403-419

Autant-Bernard, C., Billand, P., Frachisse, D., and 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

Baltagi, B. (2008): *Econometric Analysis of Panel Data*, 4th Edition, Wiley

Barber, M.J., Fischer, M.M. and Scherngell, T. (2011): The community structure of R&D cooperation in Europe: evidence from a social network perspective. *Geographical Analysis* 43, 415-432.

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

Bergman, E. and Maier, G. (2009): Network central: Regional positioning for innovative advantage. *The Annals of Regional Science* 43, 615-644

Bonacich, P. (1987): Power and centrality: A family of measures. *American Journal of Sociology* 92, 1170-82

Borgatti, S.P. (2005): Centrality and network flow, *Social Networks* 27, 55-71

Borgatti, S.P. and Everett, M.G (2006): A Graph-theoretic perspective on centrality. *Social Networks* 28, 466–484

Boschma, R. (2005): Proximity and innovation: a critical assessment. *Regional Studies* 39, 61-74

Breschi, S. and Lissoni, F. (2009): Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, 9: 439-468

Breschi, S. and Malerba, F. (2009): ERA and the role of networks. In: Delanghe, H., Muldur U. and Soete L. (eds): *European Science and Technology Policy. Towards Integration or Fragmentation*?, Cheltenham, Edward Elgar, 160-174

Broekel, T. and Brenner, T. (2011): Regional factors and innovativeness: An empirical analysis of four German industries, *The Annals of Regional Science* 47, 169-194

Cohen, W.M. and Levinthal, D.A. (1990): Absorptive capacity: A new perspective on learning and innovation, *Administrative Science Quarterly* 35, 128-152

Commission of the European Communities (CEC) (2007): Green paper 'The European Research Area: New Perspectives'. {SEC(2007) 412}, COM(2007)161 final, Brussels, 4 April 2007

Cowan, R. and Jonard, N. (2004): Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control* 28, 1557-1575.

Elhorst, J.P. (2003): Specification and estimation of spatial panel data models. *International Regional Science Review* 26(3), 244–68

Elhorst, J.P. (2012): Matlab Software for Spatial Panels. International Regional Science Review, doi: 0160017612452429

Feldman, M.P. and Kogler, D. (2010): Stylized Facts in the Geography of Innovation. In: Hall, H. and Rosenberg, N. (eds.) *Handbook of the Economics of Innovation*, Volume 1, Elsevier, North Holland, 382-410

Fischer, M.M. and Varga, A. (2003): Spatial knowledge spillovers and university research. *The Annals of Regional Science* 37, 302-322

Fischer, M.M., Scherngell, T. and Jansenberger, E. (2006): The geography of knowledge spillovers between high-technology firms in Europe. Evidence from a spatial interaction modelling perspective. *Geographical Analysis* 38, 288-309

Fischer M.M., Scherngell, T. and Reismann, M. (2009): Knowledge spillovers and total factor productivity. Evidence using a spatial panel data model. *Geographical Analysis* 41, 204-220

Freeman, L.C. (1979): Centrality in social networks: Conceptual clarification. *Social Networks* 1(3), 215-239

Gilsing, V., Nooteboom, B. Vanhaverbekec, W., Duystersd, G and van den Oorda, A. (2008): Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy* 37, 1717–1731

Goetze, C. (2010): An empirical enquiry into co-patent networks and their stars: The case of cardiac pacemaker technology. *Technovation* 30 (7/8), 436–446

Grossman, G. and Helpman, E. (1991): *Innovation and Growth in the Global Economy*. MIT Press

Hagedoorn, J., Link, A.N. and Vonortas, N.S. (2000): Research partnerships. *Research Policy* 29(4/5), 567-586

Hicks, D.M. and Katz, J.S. (1996): Where is science going? Science, Technology & Human Values 21, 379-406.

Hoekman, J., Frenken, K., and van Oort, F. (2009): The geography of collaborative knowledge production in Europe. *The Annals of Regional Science* 43, 721-738

Karlsson, C. and Manduchi, A. (2001): Knowledge Spillovers in a Spatial Context – A critical review and assessment. In: Fischer, M. M. and Fröhlich, J. (eds.): *Knowledge, Complexity and Innovation Systems*, Springer, Berlin, Heidelberg and New York, 101-124

Katz, J.S. and Martin, B.R. (1997): What is research collaboration? *Research Policy* 26, 1-18

Knoke, D. and Young, S. (2008): *Social Network Analysis.* Los Angeles, Sage

Lagendijk, A. (2001): Scaling knowledge production: How significant is the region. In: Fischer, M. M. and Fröhlich, J. (eds.): *Knowledge, Complexity and Innovation Systems*, Berlin, Heidelberg and New York, Springer, 79-100

Le Sage, J. and Pace, R.K. (2009): *Introduction to spatial econometrics*. CRC Press, Boca Raton, London and New York

Maggioni, M.A. and Uberti, T.E. (2009): Knowledge networks across Europe: which distance matters? *The Annals of Regional Science* 43, 691-720

Mairesse, J. and Turner, L. (2005): Measurement and explanation of the intensity of co-publication in scientific research: an analysis at the laboratory level. NBER Working Paper, No. 11172, Cambridge/MA.

Newman, M.E.J. (2001): Scientific collaboration networks. I. Network construction and fundamental results. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 64 (1 II), 016131/1-016131/8

Paier, M. and Scherngell, T. (2011): Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model. *Industry and Innovation* 18, 89-104

Park, H.M (2011): Practical Guides To Panel Data Modeling: A Step-by-step Analysis Using Stata. Tutorial Working Paper. Graduate School of International Relations, International University of Japan.

Ponds, R., van Oort, F. and Frenken, K. (2007): The geographical and institutional proximity of research collaboration. *Papers in Regional Science* 86 (3), 423-443

Powell, W.W. and Giannella, E. (2010): Collective Invention and Inventor Networks. In: Hall, H. and Rosenberg, N. (eds.) *Handbook of the Economics of Innovation*, Volume 1, Elsevier, North Holland, 576-605

Powell, W.W. and Grodal, S. (2005): Networks of innovators. In: Fagerberg, J., Mowery, D.C. and Nelson, R.R. (eds.): *The Oxford Handbook of Innovation*, Oxford, Oxford University Press, 56-85

Romer, P. M. (1990): Endogenous technological change. *Journal of Political Economy* 98, 71-102

Scherngell, T. and Barber, M. (2011): Distinct spatial characteristics of industrial and public research collaborations: Evidence from the 5th EU Framework Programme. *The Annals of Regional Science* 46, 247-266

Scherngell, T. and Barber, M. (2009): Spatial interaction modelling of cross-region R&D collaborations. Empirical evidence from the 5th EU Framework Programme. *Papers in Regional Science* 88, 531-546

Scherngell, T. and Lata, R. (2012): Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data. *Papers in Regional Science*. doi: 10.1111/j.1435-5957.2012.00419.x

Singh, J. (2005): Collaborative networks as determinants of knowledge diffusion patterns. *Management Science* 51(5), 756–770

Sorenson, O., Rivkin J.W. and Fleming, L. (2006): Complexity, networks and knowledge flow. *Research Policy* 35(7), 994–1017

Stephan, P.E. (2010): The Economics of Science. In: Hall, H. and Rosenberg, N. (eds.) *Handbook of the Economics of Innovation*, Volume 1, Elsevier, North Holland, 217–273

Ter Wal, A.L.J. and Boschma, R. (2009): Applying social network analysis in economic geography: framing some key analytical issues. *The Annals of Regional Science* 43, 739-756

Varga, A. (2001): Universities and regional economic development: does agglomeration matter? In: Johansson, B., Karlsson, C. and Stough, R. (eds.) Theories of Endogenous Regional Growth – Lessons for Regional Policies, Springer, Berlin, Heidelberg and New York, 345-367

Wanzenböck, I., Scherngell, T., and Lata, R. (2012): Embeddedness of European regions in EU funded R&D networks: A spatial econometric perspective. Paper presented at the first Eurolio European Seminar on Geography of Innovation, 26-28 January 2012, Congress Center, Saint-Etienne (France)

Wasserman, S. and Faust, K. (1994): *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge

Wuchty, S., Jones, B.F. and Uzzi, B. (2007): The increasing dominance of teams in the production of knowledge. *Science* 316, 1036-1039

Zucker, L.G. and Darby, M.R. (2007): Virtuous circles in science and commerce. *Papers in Regional Science* 86, 445-470

Appendix

List of regions

NUTS is an acronym of the French for the “nomenclature of territorial units for statistics", which is a hierarchical system of regions used by the statistical office of the European Community for the production of regional statistics. At the top of the hierarchy are NUTS-0 regions (countries) below which are NUTS-1 regions and then NUTS-2 regions. This study disaggregates Europe's territory into 241 NUTS-2 regions located in the EU-25 member states (except Cyprus and Malta). We exclude the Spanish North African territories of Ceuta y Melilla, the Portuguese non-continental territories Azores and Madeira, and the French Departments d'Outre-Mer Guadeloupe, Martinique, French Guayana and Reunion. Thus, we include the following NUTS 2 regions:

Austria:	Burgenland, Kärnten, Niederösterreich, Oberösterreich, Salzburg, Steiermark, Tirol, Vorarlberg, Wien
Belgium:	Prov. Antwerpen, Prov. Brabant-Wallon, Prov. Hainaut, Prov. Limburg (B), Prov. Liège, Prov. Luxembourg (B), Prov. Namur, Prov. Oost-Vlaanderen, Prov. Vlaams-Brabant, Prov. West-Vlaanderen, Région de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest
Czech Republic:	Jihovýchod, Jihozápad, Moravskoslezsko, Praha, Severovýchod, Severozápad, Střední Morava, Střední Čechy
Denmark:	Danmark
Estonia:	Eesti
Finland:	Åland, Etelä-Suomi, Itä-Suomi, Länsi-Suomi, Pohjois-Suomi
France:	Alsace, Aquitaine, Auvergne, Basse-Normandie, Bourgogne, Bretagne, Centre, Champagne-Ardenne, Corse, Franche-Comté, Haute-Normandie, Île de France, Languedoc-Roussillon, Limousin, Lorraine, Midi-Pyrénées, Nord - Pas-de-Calais, Pays de la Loire, Picardie, Poitou-Charentes, Provence-Alpes-Côte d'Azur, Rhône-Alpes
Germany:	Arnsberg, Berlin, Brandenburg, Braunschweig, Bremen, Chemnitz, Darmstadt, Dessau, Detmold, Dresden, Düsseldorf, Freiburg, Gießen, Halle, Hamburg, Hannover, Karlsruhe, Kassel, Koblenz, Köln, Leipzig, Lüneburg, Magdeburg, Mecklenburg-Vorpommern, Mittelfranken, Münster, Niederbayern, Oberbayern, Oberfranken, Oberpfalz, Rheinhessen-Pfalz, Saarland, Schleswig-Holstein, Schwaben, Stuttgart, Thüringen, Trier, Tübingen, Unterfranken, Weser-Ems
Greece:	Anatoliki Makedonia, Thraki; Attiki; Ipeiros; Voreio Aigaio; Dytiki Ellada; Dytiki Makedonia; Thessalia; Ionia Nisia; Kentriki Makedonia; Kriti; Notio Aigaio; Peloponnisos; Sterea Ellada
Hungary:	Dél-Alföld, Dél-Dunántúl, Észak-Alföld, Észak-Magyarország, Közép-Dunántúl, Közép-Magyarország, Nyugat-Dunántúl
Ireland:	Border, Midland and Western; Southern and Eastern

Italy:	Abruzzo, Basilicata, Calabria, Campania, Emilia-Romagna, Friuli-Venezia Giulia, Lazio, Liguria, Lombardia, Marche, Molise, Piemonte, Puglia, Sardegna, Sicilia, Toscana, Trentino-Alto Adige, Umbria, Valle d'Aosta/Vallée d'Aoste, Veneto
Latvia:	Latvija
Lithuania:	Lietuva
Luxembourg:	Luxembourg (Grand-Duché)
Netherlands:	Drenthe, Flevoland, Friesland, Gelderland, Groningen, Limburg (NL), Noord-Brabant, Noord-Holland, Overijssel, Utrecht, Zeeland, Zuid-Holland
Poland:	Dolnośląskie, Kujawsko-Pomorskie, Lubelskie, Lubuskie, Łódzkie, Mazowieckie, Małopolskie, Opolskie, Podkarpackie, Podlaskie, Pomorskie, Śląskie, Świętokrzyskie, Warmińsko-Mazurskie, Wielkopolskie, Zachodniopomorskie
Portugal:	Alentejo, Algarve, Centro (P), Lisboa, Norte
Slovakia:	Bratislavský kraj, Stredné Slovensko, Východné Slovensko, Západné Slovensko
Slovenia:	Slovenija
Spain:	Andalucía, Aragón, Cantabria, Castilla y León, Castilla-La Mancha, Cataluña, Comunidad Foral de Navarra, Comunidad Valenciana, Comunidad de Madrid, Extremadura, Galicia, Illes Balears, La Rioja, País Vasco, Principado de Asturias, Región de Murcia
Sweden:	Mellersta Norrland, Norra Mellansverige, Småland med öarna, Stockholm, Sydsverige, Västsverige, Östra Mellansverige, Övre Norrland
United Kingdom:	Bedfordshire & Hertfordshire; Berkshire, Buckinghamshire & Oxfordshire; Cheshire; Cornwall & Isles of Scilly; Cumbria; Derbyshire & Nottinghamshire; Devon; Dorset & Somerset; East Anglia; East Riding & North Lincolnshire; East Wales; Eastern Scotland; Essex; Gloucestershire, Wiltshire & North Somerset; Greater Manchester; Hampshire & Isle of Wight; Herefordshire, Worcestershire & Warkwickshire; Highlands and Islands; Inner London; Kent; Lancashire; Leicestershire, Rutland and Northamptonshire; Lincolnshire; Merseyside; North Eastern Scotland; North Yorkshire; Northern Ireland; Northumberland and Tyne and Wear; Outer London; Shropshire & Staffordshire; South Western Scotland; South Yorkshire; Surrey, East & West Sussex; Tees Valley & Durham; West Midlands; West Wales & The Valleys; West Yorkshire