

Integration Processes in European R&D: A comparative spatial interaction approach using project based R&D networks and co-patent networks

Rafael Lata¹, Thomas Scherngell¹ and Thomas Brenner²

¹Foresight and Policy Development Department,
Austrian Institute of Technology (AIT) , Vienna, Austria

²Economic Geography and Location Research,
Philipps University Marburg

Abstract. The focus of this study is on integration processes in European R&D by analyzing the spatio-temporal dimension of two different R&D collaboration networks across Europe. These networks cover different types of knowledge creation, namely co-patent networks and project based R&D networks within the EU Framework Programmes (FPs). Integration in European R&D – one of the main pillars of the EU Science Technology and Innovation (STI) policy – refers to the harmonisation of fragmented national research systems across Europe and to the free movement of knowledge and researchers. The objective is to describe and compare spatio-temporal patterns at a regional level, and to estimate the evolution of separation effects over the time period 1999-2006 that influence the probability of cross-region collaborations in the distinct networks under consideration. The study adopts a spatial interaction modeling perspective, econometrically specifying a panel generalized linear model relationship, taking into account spatial autocorrelation among flows by using Eigenfunction spatial filtering methods. The results show that geographical factors are a lower hurdle for R&D collaborations in FP networks than in co-patent networks. Further it is shown that the geographical dynamics of progress towards more integration is higher in the FP network.

JEL Classification: O38, L14, R15

Keywords: R&D networks, European Framework Programme, European Research Area, Spatial interaction modelling, Eigenvector spatial filtering

1 Introduction

Today it is widely recognised that *first*, innovation processes are increasingly based on interaction, research collaborations and networks of various actors (see, for instance, Fischer 2001, Powell and Grodal 2005)¹, and, *second*, innovation is the key element for sustained economic growth of firms, industries, regions and countries (see, for example, Romer 1990)². Based on these arguments, the main focus of the Europe 2020 Strategy is on research and innovation in order to achieve a new growth path that leads to a smart, sustainable and inclusive economy (European Commission 2011). In this context, the concept of the *Innovation Union* – one of the seven flagships scheduled in the Europe 2020 Strategy – is intended to improve conditions for innovation and knowledge diffusion to ensure that innovative ideas are efficiently turned into new products and services that create growth and employment (European Commission 2010). One of the main pillars of the *Innovation Union* is the realisation of an integrated European Research Area (ERA), defined as one explicit principal purpose to fulfil progress towards the Innovation Union.

The concept of the European Research Area (ERA) refers to the objective to enable and facilitate “free circulation of researchers, knowledge and technology” across the countries of the EU, and, by this, stimulating integration processes in European R&D (see Commission of the European Union (CEU) 2008, p. 6). This policy goal is to be addressed by improving coherence of the European research landscape, thus removing barriers – such as geographical, cultural, institutional and technological impediments – for knowledge flows, knowledge diffusion and researcher mobility by a European-wide coordination of national and regional

¹ The literature on R&D networks underlines the crucial importance of cooperative agreements between universities, companies and governmental institutes, for developing and integrating new knowledge in the innovation process (see Powell and Grodal 2005). This is explained by considerations that innovation nowadays takes place in an environment characterised by uncertainty, increasing complexity and rapidly changing demand patterns in a globalised economy. Organisations must collaborate more actively and more purposefully with each other in order to cope with increasing market pressures in a globalizing world, new technologies and changing patterns of demand. In particular, firms have expanded their knowledge bases into a wider range of technologies (Granstrand 1998), which increases the need for more different types of knowledge, so firms must learn how to integrate new knowledge into existing products or production processes (Cowan 2004). It may be difficult to develop this knowledge alone or acquire it via the market. Thus, firms form different kinds of co-operative arrangements with other firms, universities or research organisations that already have this knowledge to get faster access to it.

² The theory of endogenous growth, and the geography-growth synthesis both consider that economic growth and spatial concentration of economic activities emanate from localised knowledge diffusion processes, in particular transferred via network arrangements between different actors of the innovation system (see, for instance, Lucas 1988, Grossman and Helpman 1991).

research activities and policy programmes, including a considerable amount of jointly-programmed public research investment (see Delanghe et al. 2009).

To gain insight into the nature of integration processes in European R&D, there is urgent need for analysing the geographical dimension of R&D networks across Europe from a longitudinal perspective. The geography of such networks has – from a static perspective – attracted increasing interest in Regional Science and Economic Geography in the recent past. While from its beginning, the measurement of such phenomena has faced numerous problems³, the empirical investigation of knowledge flows and R&D collaborations has significantly improved during the 1990s by using new indicators such as patent citations (see, for instance, the pioneering study by Jaffe et al. 1993, Fischer et al. 2006), co-publications (see, for instance, Ponds et al. 2009, Hoekman et al. 2010) or project based R&D networks within the FPs (see Scherngell and Barber 2009 and 2011), and by introducing new methods, in particular new spatial econometric techniques (see, for instance, Anselin et al. 1997, LeSage et al. 2007). Recent studies focus on the structure of knowledge flows by adopting a spatial interaction modelling perspective (see, for instance, Ponds et al. 2007, Scherngell and Barber 2009), employing a social network analysis perspective (see, for instance, Breschi and Cusmano 2004, Barber et al. 2011), or a combination of both (see Barber and Scherngell 2011).

However, as these studies just provide a static picture on the geography of R&D collaborations, novel questions arise – both in theoretical and empirical research as well as in a European policy context – regarding R&D network structures and its dynamics. Concerning the particular focus on integration processes in European R&D, the evolution of different kinds of separation effects over time – such as geographical, technological, institutional or cultural barriers – that determine R&D collaboration networks is of crucial interest. Thus, this study shifts emphasis to the investigation of the geographical dynamics of two different types of R&D collaboration networks across Europe, namely co-patent networks and project based R&D networks within the European Framework Programmes (FPs). We take these types of R&D collaboration networks to analyse integration processes in European R&D over time

³ “Knowledge flows are invisible, they leave no paper trail by which they may be measured and tracked” states Krugman (1991, pp. 153), pointing to difficulties in finding data on knowledge diffusion with respect to the geographical and temporal location of such processes.

from two different angles, shifting attention to a comparison of European integration processes in these networks.

By this, the study addresses one of the major drawbacks of the current empirical literature: the lack of a longitudinal and comparative perspective of distinct R&D collaboration networks. Some exceptions are the studies of Maggioni and Uberti (2009), Hoekman et al. (2010), Hoekman et al. (2012) and Scherngell and Lata (2012). Hoekman et al. (2010) and Scherngell and Lata (2012) investigate the ongoing process of European integration by determining the impact of geographical distance and territorial borders on the probability of research collaborations between European regions. By analysing co-publication and FP network patterns and trends, the authors show that geographical distance has a negative effect on co-publication activities and FP cooperation, while for the FP networks this effect decreases over time. The study of Maggioni and Uberti (2009) focuses on the structure of knowledge flows by analysing four distinct collaboration networks, including co-patenting. Hoekman et al. (2012) focus on the effect of participation in FP networks on subsequent international publications, showing that the FPs indeed positively influence international co-publications, and, by this, seem to enhance integration across Europeans research systems.

The current study intends to complement the picture drawn in these studies, by shifting attention to a longitudinal and comparative perspective on two different R&D networks across Europe. The objective is to identify and compare the evolution of geographical, technological, institutional or cultural effects that influence the probability for collaboration activities in the different collaboration networks, and provide direct evidence on integration processes in European R&D from different angles. We adopt a regional perspective that is an appropriate approach to observe different R&D collaboration networks in geographical space (see, for instance, Hoekman et al. 2010, Scherngell and Barber 2009) over the time period 1999-2006. The study employs a Poisson spatial interaction modelling perspective to address these research questions. We adjust the spatial interaction models by accounting for spatial autocorrelation issues of flows by means of Eigenvector spatial filtering (see Chun 2008, Scherngell and Lata 2012).

The paper is organised as follows. Section 2 sets forth the conceptual background of the study with a special focus on ERA, before Section 3 reflects on the different types of R&D collaboration networks under consideration. Section 4 describes the empirical setting and the

data, accompanied by some descriptive statistics and exploratory spatial data analysis. Section 5 specifies the empirical model in form of a panel version of the spatial interaction modelling framework that is used to identify the evolution of separation effects influencing the probability of cross-region collaboration activities in the distinct networks. Section 6 presents the modelling results, before Section 7 closes with a summary of the main results and some conclusions in a European policy context.

2 The ERA goal of progress towards more integration in European R&D

One significant turning point in the EU Science, Technology and Innovation (STI) policy was the design of the concept of the European Research Area (ERA) presented at the Lisbon Council in the year 2000, rooted in the increasing awareness that European research activities suffer from diverse and fragmented national research systems (Boyer 2009). The overall goal of ERA is to overcome fragmentation in the European research system and to address the establishment of an ‘internal market’ for research across Europe, where researchers, technology and knowledge are supposed to circulate freely (see Delanghe et al. 2009, European Council 2000). The ERA green paper (CEC 2007) underlines the overall objectives of the Lisbon strategy, emphasising that the future European science and research landscape should be characterised by an adequate flow of competent researchers with high levels of mobility between institutions by integrated and networked research infrastructures and effective knowledge sharing, notably between public research and industry. This requires the reduction of geographical, cultural, institutional, and technological obstacles in order to generate research collaboration across European regions and countries (see, for instance, Breschi and Malerba 2009, Hoekman et al. 2012, Scherngell and Lata 2012).

The Framework Programmes (FPs) of the European Commission (EC) constitute the main instrument to achieve this goal, shifting emphasis on supporting and stimulating collaborative R&D activities between innovating organisations across Europe, in particular firms and universities. At the same time, regional and national research policies deal with similar issues as reflected by a growing awareness among national policy makers that national efforts are often insufficient to keep pace in the international innovation competition. In this context the European Council underlined the importance of cross border cooperation for the achievement of these objectives and put collaborative R&D activities at the centre of its strategy (Guzzetti

2009). Svanveldt (2009) highlights the crucial importance of cross-border cooperation as instrument for adequately dealing these challenges.

During the last decade, the ERA concept has been developed further (Luukkonen 2010), becoming strong political support in the context of the conception of the so-called *Innovation Union* (European Commission 2010). As one of the seven flagships scheduled in the *Europe 2020 Strategy*, the Innovation Union is intended to improve framework conditions for innovation and knowledge diffusion. Moreover one of the main objectives of the Innovation Union is to “... quickly taking all measures necessary for a well functioning and coherent European Research Area in which researchers, scientific knowledge and technology circulate freely, in which RDI investments are less fragmented and the intellectual capital across Europe can be fully exploited” (European Commission 2010, p.7). In order to tackle these challenges, specific commitments have been introduced. One of these commitments is to complete the ERA by 2014 with the goal to remove the remaining obstacles for collaborative knowledge production and consequently to foster the integration in the European research landscape (European Commission 2010).

With this in mind, the present study aims to evaluate the progress towards more integration in European R&D – as formulated in the concept of ERA and the Innovation Union. To gain empirical insight into the nature of such integration processes across Europe, the study focuses on a broad spectrum of R&D collaboration activities, namely co-patent networks and project based R&D networks within the FPs. In estimating the evolution of separation effects that capture the above mentioned obstacles for collaborative knowledge production across Europe, the analysis will show distinct mechanisms of integration processes corresponding to the different types of R&D networks. The section that follows reflects on the two different networks types under consideration in some detail.

3 A Network perspective on Integration in European R&D

R&D networks – defined as sets of organisations performing joint R&D activities – attracted burst of attention in the recent past as essential element of modern knowledge production and innovation processes (see, for instance, Castells 1996). In the current study, we take such network arrangements across Europe to analyse integration processes in European R&D,

focusing on two different types of networks that capture different types of knowledge production processes. We focus on R&D networks in the form of joint patenting, resulting in co-patents and project based R&D networks within the FPs.

Co-patent networks mainly reflect research collaborations that are related to applied knowledge generation (see, for instance, Makri et al. 2011) focusing on the development of marketable innovations and industry research activities (Maggioni et.al 1999, Götze 2011). Patents represent a well established indicator of knowledge generation activities and are widely used in empirical studies on knowledge flows (see, for instance, Jaffe et al. 1993, Fischer et. al 2006). A co-patent is defined as a patent invented by at least two inventors from two different organisations. Therefore, it represents knowledge exchange across actors within an inventor network in the process of patenting an invention (see, for instance, Ejermo and Karlsson 2006, Ter Wal and Boschma 2008, Goetze 2011).

The second type of the R&D networks refers to project based R&D collaboration within the FPs. While co-patent networks mainly reflect applied research, project based FP networks involve basic and applied research aspects, given by the fact that publications and patents may be outputs of FP networks. In the FP network, the research collaboration is constituted by joint R&D projects conducted by organisations distributed across Europe. The FPs are the main political instrument to support pre-competitive collaborative R&D within the European Union. The key objectives are, *first*, to strengthen the scientific research and technological development in the scientific landscape and by this to foster the European competitiveness, and, *second*, to promote research activities in support of other EU policies (Maggioni et.al 2009)⁴.

FP projects share specific characteristics (see for example Roediger-Schluga and Barber 2006, Breschi and Malerba 2009). *First*, they are all promoted by self-organised consortia and have distinct partners – for instance individuals, industrial and commercial firms, universities, research organisations, etc. – that are located in different EU members and associated states. *Second*, they focus primarily on pre-competitive R&D projects. *Third*, they are characterised by less market orientation and longer development periods (Polt et. al 2008).

⁴ Since their introduction in 1984, different thematic aspects and issues of the European scientific landscape have been addressed by the FPs (see Breschi and Malerba 2009). Although the FPs have undergone different changes in their orientation during the past years, their fundamental rationale remained unchanged (Roediger-Schluga and Barber 2006).

Given the properties of the two different network types under consideration, it may be hypothesised that integration processes for these network types differ. This may, on the one hand, be related to the different knowledge generation processes in these networks, on the other hand, to governance rules and policy programmes implemented by the EC influencing the resulting network structures. The spatial interaction models (see Section 5) will enable us to proof this hypothesis, and disclose distinct spatial characteristics and collaboration patterns in the networks under consideration, and, by this, drawing a more detailed picture on integration processes in European R&D.

4 Data and descriptive statistics

In our empirical analysis we aim to investigate integrations processes in European R&D networks focusing on two different types of collaboration networks, that is FP collaboration networks and co-patent networks. The EUPRO database is used to capture project based R&D networks within the FPs, while the Regpat database is taken to construct co-patent networks. The *EUPRO database* currently comprises information on more than 60,000 research projects funded by the EU FPs and all participating organisations. A network link is given between two organisations when they conduct a joint research project in the FPs. We use information on the geographical location in form of the city to trace the geographical dimension of the network⁵. The *Regpat database* contains information on patent applications from various patent offices worldwide. It is provided by the OECD and contains, among many others, all patent applications issued at the European Patent Office (EPO), and the national patent offices of the European countries. A network link between two organisations is given when inventors from two different organisations appear on a patent application. We use information on the inventor address of an EPO patent application to trace the origin of the invention.

The European coverage is achieved by using $i, j = 1, \dots, n$ NUTS-2 regions⁶ of the 25 pre-2007 EU member-states as well as Norway and Switzerland. We extract n -by- n collaboration

⁵ EUPRO has been constructed by AIT by substantially standardising raw data on EU FP research collaborations obtained from the CORDIS database. Standardisation included identification of unique organisation names (i.e. entries are assigned to the respective unique organisation name), identification of unique organisation types, etc (see Roediger-Schluga and Barber 2006 for details on the database).

⁶ Although substantial size differences and interregional disparities of some regions exist, these units are widely recognized to be an appropriate level for modelling and analysis purposes (see, for example, LeSage and Fischer 2012).

matrices for each time period $t = 1, \dots, T$, both for the FP- and for the co-patent network, by aggregating the number of individual collaborative activities at the organisational level in time period t to the regional level. This leads to the observed number of R&D collaborations y_{ijt} between two regions i and j in time period t in the respective network, that is the FP and the co-patent network. The resulting regional collaboration matrix Y_t for the two networks⁷ for a given year t contains the collaboration intensities between all (i, j) -region pairs, given the $i = 1, \dots, n$ regions in the rows and the $j = 1, \dots, n$ regions in the columns. We use a full counting procedure⁸ for the construction of our collaboration matrices (see, for example, Katz 1994).

Figure 1 illustrates the spatial distribution of the cross-region R&D collaborations in the FP- (Figure 1a) and the co-patent network (Figure 1b) across Europe. In the spatial network maps, the sizes of the nodes are proportional to the number of regional participations in the two distinct networks. The darkness of the lines corresponds with the number of joint R&D collaborations between two regions, i.e. the darker the higher the interaction intensity. It is shown that the spatial structures of the distinct networks differ markedly. The most striking difference concerns the fact that the international collaboration activity is much higher in the FP network than in the co-patent network. In the latter, R&D collaborations are widely confined within national boundaries, while such boundaries seem to play a minor role for the structure of the FP network. Furthermore, the intra-regional collaboration intensity seems much higher in the co-patent network than in the FP network, pointing to the geographical localisation of the co-patents within NUTS-2 regions, while the cross-region collaboration intensity is much higher in the FP network. Concerning the spatial distribution of the regions with high intra-regional co-patent activities, a high intensity can be found for regions belonging to the traditional industrial core of Europe (see Hoekman et al. 2012), also referred to as the European ‘blue banana’ (Brunet 2002), while the participation within the FP network seems to be spatially more dispersed.

⁷ Note that we do not distinguish between the FP network and the co-patent in the formal description of data as well as the modelling approach in the section that follows.

⁸ For a project with, for example, three different participating organizations a, b and c, which are located in three different regions, we count three links (from a to b, from b to c and from a to c).

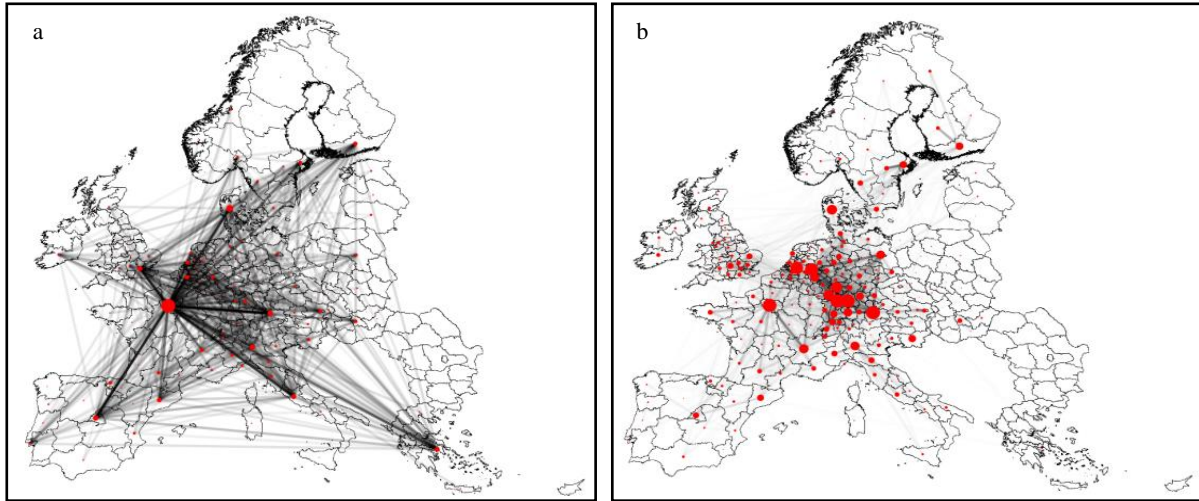


Figure 1: Spatial distribution of the cross-region R&D networks for the year 2006. (a): R&D collaborations within the FP-network. (b): R&D collaborations within the co-patent network.

However, both networks seem to be spatially concentrated in some European regions that show high collaboration intensity. In this context the question arises, whether a spatial clustering of interaction patterns in the two networks can be observed, and which network shows a higher degree of spatial clustering, also referred to as spatial autocorrelation of flows (see, for instance, Berglund and Karlstrom 1999). Spatial autocorrelation of flows is, for example, when flows from a particular origin may be correlated with other flows that have the same origin, and, similarly, flows into a particular destination may be correlated with other flows that have the same destination (Scherngell and Lata 2012). In our case, this means that the intensity of R&D collaborations from an origin region i to a destination region j may be correlated with the intensity of R&D collaborations from the same origin i to another destination j , or vice versa. Such a situation is both interesting from the perspective of our research question on integration in European R&D, namely by assessing whether such R&D collaborations are statistically concentrated to a geographical core of regions that are located nearby to each other. From a theoretical perspective the spatial autocorrelation of R&D collaboration flows may be explained by the assumption that the collaboration behaviour of one region influences the collaboration behaviour of neighbouring regions because – as described in various empirical studies – contiguity of regions may induce knowledge flows between them, to them, and from them, and, thus, evoke the transfer of information on potential collaboration partners that are located further away (Scherngell and Lata 2012)⁹.

⁹ To give an example, if region A has many collaborations with region B (that is no neighbour of region A), region A may influence a neighbouring region C also to collaborate with region B due to information flows between region A and region C, in particular flows of ‘know who’ type information (see Cohen and Levinthal 1990).

In order to test for the existence of spatial autocorrelation of flows, we calculate a Moran's I test for spatial dependence as widely used in exploratory spatial data analysis (see, for instance, Griffith 2011). For our empirical analysis, the Moran's I_t test for a time period t is calculated as by

$$I_t = \frac{\mathbf{y}_t' \mathbf{W}^* \mathbf{y}_t}{\mathbf{y}_t' \mathbf{y}_t} \quad (1)$$

where \mathbf{y}_t is a vector of our observed collaboration flows at time t with $N=n^2$ elements $(y_{ijt}) = (y_{11t}, \dots, y_{1nt}, y_{21t}, \dots, y_{2nt}, \dots, y_{n1t}, \dots, y_{nnt})$, and \mathbf{W}^* is defined by $\mathbf{W} \otimes \mathbf{W}$ where \mathbf{W} is the n -by- n spatial weights matrix and \otimes denotes the Kronecker product. For \mathbf{W} , we set

$$w_{ij} = \begin{cases} 1 & \text{if } s_{ij}^{(1)} \leq s_{ig(i)}^{(1)} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where $s_{ij}^{(1)}$ measures the great circle distance between the economic centers of two regions i and j , and $g(i)$ denotes the g -nearest neighbour of i . We define $g = 5$, as used in various empirical studies dealing with European regions (see, for instance, Parent and LeSage 2012). The respective Moran's I statistics for the years 1999-2006 are reported in Table 1. The results are most often significant pointing to substantial spatial autocorrelation of R&D collaborations in both networks under consideration, i.e. a high number of flows is correlated with flows that come from nearby origins, and going into nearby destinations. However, the degree of spatial dependence is much higher for the co-patent network as has been expected considering the spatial distribution of the flows that are visualised in Figure 1. Furthermore, the Moran's I for the FP network shows a decreasing trend, while for the co-patent network no time trend can be observed, pointing to differences in integration processes for the two network types. This is one particular point to be tested in the empirical model that is introduced in the section that follows, disclosing the influence of different exogenous variables that may explain these spatial collaboration patterns. In this context, the existence of spatial autocorrelation also bears important implications in a modeling context, since estimates may be biased neglecting spatial autocorrelation issues of flows (see, for instance, Fischer and Griffith 2008, LeSage and Pace 2008, Scherngell and Lata 2012)

Table1. Spatial autocorrelation of R&D collaboration in two distinct networks (1999-2006)

	Moran' <i>I</i>							
	1999	2000	2001	2002	2003	2004	2005	2006
<i>FP- Network</i>	0.016*	0.006*	0.003*	0.000	-0.001	0.007*	-0.009	-0.001
<i>Co- Patent Network</i>	0.136*	0.120*	0.132*	0.144*	0.139*	0.153*	0.146*	0.147*

*significant at the 0.001 significance level,

5 The empirical model

This section shifts direct attention to the modelling approach used to address the basic research question of the study that is to estimate how specific separation effects influence the variation of cross-region R&D collaborations in two distinct collaboration networks over time, and, by this, providing direct evidence on distinct integration processes in different types of R&D. To address this question, we employ a spatial interaction modelling approach¹⁰. In implementing a panel version of the spatial interaction model, we are able to identify time effects that are necessary to observe potential integration processes of the networks over the time period 1999-2006. In what follows we will specify the panel version of the spatial interaction model, an extension accounting for spatial autocorrelation issues of flows, and describe the independent variables of the model.

The panel version of the spatial interaction model to be estimated

Let us denote Y_{ijt} as a random dependent variable corresponding to observed R&D collaborations y_{ijt} within the FP- or the co-patent network between origin i ($i = 1, \dots, n$) and destination j ($j = 1, \dots, n$) at time t ($t = 1, \dots, T$). As in the previous section, we do not distinguish between the two networks in the formal model presentation; our basic model is given by

$$Y_{ijt} | y_{ijt} = \mu_{ijt} + \varepsilon_{ijt} \quad i, j = 1, \dots, n; \quad t = 1, \dots, T \quad (3)$$

¹⁰ Spatial interaction models are widely used for modelling origin-destination flows data and were used to explain different kinds of flows, such as migration, transport or communication flows, between discrete units in geographical space (see, for instance, Sen and Smith 1995, LeSage and Fischer 2010 among many others).

where μ_{ijt} denotes some mean expected interaction frequency between origin i and destination j at time t , ε_{ijt} some disturbance term about the mean with the property $E[\varepsilon_{ijt} | y_{ijt}] = 0$. As in classical spatial interaction theory (see, for instance, Fischer and Wang 2011), we model the mean interaction frequencies μ_{ijt} between origin i and destination j at time t by some origin function O_{it} which characterizes the origin i of interaction in time period t , some destination function D_{jt} which describes the destination j of interaction in time period t , and some separation function S_{ijt} which accounts for the separation between an origin region i and a destination region j in time period t . Then we use a multiplicative relationship for our basic model, given by

$$\mu_{ijt} = O_{it} D_{jt} S_{ijt} \quad i, j = 1, \dots, n; \quad t = 1, \dots, T \quad (4)$$

where

$$O_{it} = o_{it}^{\alpha_1} \quad i, j = 1, \dots, n; \quad t = 1, \dots, T \quad (5)$$

$$D_{jt} = d_{jt}^{\alpha_2} \quad i, j = 1, \dots, n; \quad t = 1, \dots, T \quad (6)$$

$$S_{ijt} = \exp \left[\sum_{k=1}^K \beta_k s_{ijt}^{(k)} \right]. \quad i, j = 1, \dots, n; \quad t = 1, \dots, T \quad (7)$$

o_{it} and d_{jt} are origin and destination variables, $s_{ijt}^{(k)}$ are K ($k = 1, \dots, K$) separation variables that are introduced below. α_1 , α_2 and β_k are parameters to be estimated.

As has come into fairly wide use for spatial interaction models, we assume $(Y_{ij}) \sim \text{Poisson}$ due to the true integer non-negative count nature of our R&D collaboration flows (see, for instance, Cameron and Trivedi 1998, Long and Freese 2001, Fischer et al. 2006, Scherngell and Barber 2009). The resulting panel version of the Poisson spatial interaction model is given by,

$$\mu_{ijt} = \exp \left[\alpha_1 \log(o_{it}) + \alpha_2 \log(d_{jt}) + \sum_{k=1}^K \beta_k s_{ijt}^{(k)} + \gamma_{ij} \right] \quad (8)$$

where γ_{ij} denotes the unobserved individual specific effect, also referred to as the one-way error component model (see Baltagi 2008). The random term γ_{ij} is time invariant but varies across all (i, j) -region pairs. In our case γ_{ij} accounts for region-pair specific effects that are

not included in the model. We assume the γ_{ij} to be correlated across our time periods for the same

(i, j) -region pair, i.e. we follow a *random effects* specification, and integrate out the random effect γ_{ij} of the joint probability $\prod_{t=1}^T \Pr(y_{ij1}, \dots, y_{ijt})$ by obtaining

$$\Pr(y_{ij1}, \dots, y_{ijt}) = \int \Pr(y_{ij1}, \dots, y_{ijt}, \gamma_{ij}) d\gamma_{ij} = \int \Pr(y_{ij1}, \dots, y_{ijt} | \gamma_{ij}) g(\gamma_{ij}) d\gamma_{ij}. \quad (9)$$

Note that this is the same approach used in models for event counts to condition the heterogeneity out of the Poisson model to produce the Negative Binomial model (see Baltagi 2008), i.e. when $(Y_{ij}) \sim \text{Poisson}$ with mean μ_{ijt} as given by Equation (8), and $\exp(\gamma_{ij}) \sim \text{Gamma}$, then our *random effects Negative Binomial spatial interaction model* to be estimated is

$$\Pr(y_{ij1}, \dots, y_{ijt}) = \frac{\left(\prod_{t=1}^T \mu_{ijt}^{y_{ijt}}\right) \Gamma\left(\theta + \sum_{t=1}^T y_{ijt}\right)}{\left(\Gamma(\theta) \prod_{t=1}^T y_{ijt}!\right) \left[\left(\sum_{t=1}^T \mu_{ijt}\right)^{\sum_{t=1}^T y_{ijt}}\right]} Q_i (1 - Q_i)^{\sum_{t=1}^T y_{ijt}} \quad (10)$$

with

$$Q_i = \frac{\theta}{\theta + \sum_{t=1}^T \mu_{ijt}} \quad (11)$$

where $\Gamma(\cdot)$ denotes the Gamma distribution and θ its variance. Parameter estimation is achieved via maximum likelihood estimation procedures (see Cameron and Trivedi 1998).

Accounting for spatial autocorrelation and time effects

Given the results of the spatial autocorrelation analysis of the previous section, it can be assumed that spatial dependence among our collaboration flows may lead to biased estimates. Thus, we re-specify our panel version of the Negative Binomial spatial interaction model by accounting for spatial autocorrelation issues as well as by introducing time effects enabling us to infer on time trends concerning the evolution of collaboration patterns in the two networks.

As noted by Chun (2008), maximum likelihood estimation assumes that all observations, in our case collaboration flows in our two networks under consideration, are mutually independent. A violation of this assumption may be in particular induced by spatial autocorrelation of flows leading to incorrect inferences due to inconsistency of the standard errors, and, thus, unrealistic significances (Chun 2008, Griffith 2009)¹¹. We follow Scherngell and Lata (2012) who apply a spatial filtering method to filter out spatial autocorrelation of residual flows in a Negative Binomial spatial interaction context. The essence of the spatial filtering approach is to extract eigenvectors from a modified spatial weights matrix that serve as spatial surrogates for omitted spatially autocorrelated origin and destination variables (see Griffith 2009 and 2011). These proxy variables are extracted as n eigenvectors¹² from the modified spatial weights matrix of the form $(\mathbf{I} - \mathbf{I}\mathbf{I}^T \frac{1}{n}) \mathbf{W} (\mathbf{I} - \mathbf{I}\mathbf{I}^T \frac{1}{n})$ with \mathbf{I} denoting the n -by- n identity matrix, \mathbf{I} is an n -by-1 vector of one's \mathbf{I}^T its transpose, and \mathbf{W} the n -by- n spatial weights matrix, as defined by Equation (2). The eigenvectors can be interpreted as synthetic map variables that represent specific natures and degrees of potential spatial autocorrelation (Chun 2008, Griffith 2003).

As noted by Chun and Griffith (2011) it is not appropriate to use the full set of \mathbf{E}_n eigenvectors for the construction of the spatial filter variables. Further, we face a situation where Eigenvectors have to be selected for each time period due to the panel version of the spatial interaction model (Chun 2011, Patuelli et al. 2011). We follow Patuelli et al. (2011) and select in a *first* step a subset of distinguished eigenvectors on the basis of their Moran's I values. We follow Fischer and Griffith (2008) and extract those Eigenvectors \mathbf{E}_m that show a higher Moran's I value than 0.25. In a *second* step, it is necessary to adapt these Eigenvectors to our spatial interaction framework; origin candidate eigenvectors are drawn from $\mathbf{I} \otimes \mathbf{E}_m$

¹¹ One way to capture spatial autocorrelation of flows is the use of spatial autoregressive techniques (LeSage and Pace 2009). An alternative approach is the use of spatial filtering methods. The key advantage of the spatial filtering approach is that it can be applied to any functional form and thus, does not depend on normality assumption (Patuelli et al. 2011). Consequently, we prefer the spatial filtering approach over spatial autoregressive model as we are dealing with a Poisson spatial interaction framework.

¹² The extracted eigenvectors, labelled as \mathbf{E}_n , have several characteristics. *First*, as shown by Griffith (1996), each extracted eigenvector relates to a distinct map pattern that has a certain degree of spatial autocorrelation. *Second*, the selected eigenvectors are centered at zero due to the pre and post multiplication of \mathbf{W} by the standard projection Matrix $(\mathbf{I} - \mathbf{I}\mathbf{I}^T \frac{1}{n})$. *Third*, the modification of \mathbf{W} ensures that the eigenvectors provide mutually orthogonal and uncorrelated map patterns ranging from the highest possible degree of positive spatial correlation to highest possible degree of negative spatial correlation as given by the Moran's I (MI). (Griffith 2003, Chun, and Griffith 2011). Hence, the first extracted eigenvector \mathbf{E}_1 is the one showing the highest degree of positive spatial autocorrelation that that can be achieved by any spatial recombination; the second eigenvector \mathbf{E}_2 has the largest achievable degree of spatial autocorrelation by any set that is uncorrelated with \mathbf{E}_1 until the last extracted eigenvector \mathbf{E}_n will maximize negative spatial autocorrelation (Griffith 2003).

and the destination candidate eigenvectors are obtained from $E_m \otimes I$. In a *third* step, these Eigenvectors are added as explanatory variables to $T=9$ cross-section versions of the Negative Binomial spatial interaction model, from which statistically significant Eigenvectors are identified. In a *fourth* step, we determine those eigenvectors that are significant over all time periods and define the resulting set of common origin and destination eigenvectors, E_q and E_r , respectively, as our time invariant spatial filter¹³. The time invariant spatial filter covers the total number of space-time observations, and account for spatial dependence of flows in our origin and destination data.

We add the selected origin filters E_q and destination filters E_r as regressors to our panel version of the Negative Binomial spatial interaction model. Further we introduce the subset of Z_t time dummies in order to capture aggregate year effects (Woodridge 2008)¹⁴. This leads to the spatially filtered panel version of the Negative Binomial spatial interaction model accounting for time effects, given by re-specifying the conditional mean μ_{ijt} so that

$$\mu_{ijt} = \exp \left[\sum_{q=1}^Q E_q \psi_q + \alpha_1 \log(o_{it}) + \sum_{r=1}^R E_r \varphi_r + \alpha_2 \log(d_{jt}) + \sum_{k=1}^K \beta_k s_{ijt}^{(k)} + \sum_{t=1}^T Z_t V_t + \gamma_{ij} \right] \quad (12)$$

The coefficients to be estimated for the spatial filters are ψ_q and φ_r , V_t is the associated parameter for the time dummy at time t .

Independent variables

We use one origin measure, and one destination measure for the FP network model and the co-patent network model.. For the model on the FP networks, the origin variable o_{it} is measured in terms of organizations participating in joint FP projects in region i , while the destination variable d_{jt} denotes the number of organizations participating in joint FP projects

¹³ We use an time invariant specification of the spatial filter as we assume an time invariant underlying spatial process.

¹⁴ In order to determinate changes of our separation variables we include interaction terms (see, for an overview, Wooldridge 2008). In this procedure, variables of interest, for example R&D (see, Griliches 1984), interact with time dummy variables and illustrate if effects changed over a certain time period or not. In our case (time) interaction terms represent the interaction between our separation variables and the time dummies and determinate how separation effects have changed over time. These interaction terms pick up the inter-temporal variation of our separation effect and remain only cross-sectional variation.

in region j . For the co-patent network model, the origin variable o_{it} is measured in terms of the number of co-patents in region i , while the destination variable d_{jt} denotes the number of co-patents in region j .

From the background of our research focus our interest is on $K=6$ separation measures $s_{ijt}^{(1)}$ measures the geographical distance between the economic centres of two regions i and j in time period t , by using the great circle distance¹⁵. $s_{ijt}^{(2)}$ is a neighbouring region dummy variable that takes a value of one if the regions i and j in time period t are direct neighbours, and zero otherwise. $s_{ijt}^{(3)}$ is a country border dummy variable that we use as a proxy for institutional barriers. The variable takes a value of zero if two regions i and j in time period t are located in the same country, and one otherwise. $s_{ijt}^{(4)}$ is a language dummy variable – accounting for cultural barriers that takes a value of zero if two regions i and j in time period t are located in the same language area, and one otherwise¹⁶. $s_{ijt}^{(5)}$ captures technological distance by using regional patent data from the European Patent Office (EPO). The application date is used to extract the data for each year of our time frame. We follow Moreno, Paci and Usai (2005) and construct a vector for each region i that contains region i 's share of patenting in each of the technological subclasses of the International Patent Classification (IPC)¹⁷. Technological proximity between two regions i and j in time period t is given by the uncentred correlation between their technological vectors. $s_{ijt}^{(6)}$ captures core-region effects; it is defined as a dummy variable that is set to one if two regions i and j in time period t are the location of the capital city of a country, and zero otherwise (see Scherngell and Lata 2012).

6 Estimation results

Table 2 reports the results from the estimation of the spatially filtered random effects Negative Binomial spatial interaction models as specified in the previous section. Standard

¹⁵ Note further that according to Bröcker (1989), we calculate the intraregional distance as $s_{ii}^{(1)} = (2/3)(A_i/\pi)^{0.5}$, where A_i denotes the area of region i , i.e. the intraregional distance is two third the radius of an presumed circular area.

¹⁶ Language areas are defined by the region's dominant language. However, in most cases the language areas are combined countries, as for instance Austria, Germany and Switzerland (one exception is Belgium, where the French speaking regions are separated from the Flemish speaking regions).

¹⁷ Technological subclasses correspond to the third-digit level of the IPC systems.

errors are given in brackets. The first column presents the results for the FP network, while the second column contains the estimates for the co-patent network. As can be seen, the estimates for the origin, destination and separation variables are most often statistically significant. The bottom of the table presents some model diagnostics that are of methodological interest. The dispersion parameter is statistically significant in both model versions, indicating that the Negative Binomial specification is essential to account for overdispersion in the data. A likelihood ratio test which compares the panel estimator with the pooled estimator confirms the appropriateness of the random effects specification.

The results are interesting in the context of the geography of innovation literature, but also very relevant and insightful from a European STI policy perspective. Geographical distance, as evidenced by the estimate of β_1 , exerts in both networks, the FP network and the co-patent network, a negative effect on collaboration probability, i.e. in both networks R&D collaboration intensity between two regions significantly decreases when they are located further away geographical distance, and this effect seems only to differ slightly in magnitude. However, concerning other geographical factors, we find a much stronger negative effect in the co-patent network than in the FP network. One striking result concerns the high negative effect of country borders, as evidenced by the estimate for β_3 , for the co-patent network as compared to the FP network showing that for R&D collaborations in the FPs country borders constitute only a low hurdle. In addition, co-patent networks seem to be to a high degree focused on neighbouring regions, i.e. the collaboration significantly increases when two organisations are located in regions that share a common border (β_2). This effect is much higher than in the FP network, pointing to a stronger spatial concentration and geographical localisation of R&D collaborations reflected by co-patents.

Concerning language area effects (β_4), we also find considerable differences between the FP network and the co-patent network. The negative effect of language is much higher for the co-patent network than for the FP network, i.e. the probability that organisations located in two different language areas collaborate is much lower in the co-patent network. This may be explained by the fact that the co-patent networks are much more subject to the industry sector, where such language barriers may – as suggested by results provided by Scherngell and Barber (2011) – constitute a lower hurdle than for research including public research organisations, in particular universities.

Table 2. Estimation results of the spatially filtered random effects Negative Binomial spatial interaction models

	FP-network	Co-patent network
Origin and destination variable [α_1] = [α_2]	0.955 ^{***} (0.002)	0.354 ^{***} (0.003)
Geographical distance [β_1]	-0.210 ^{***} (0.005)	-0.266 ^{***} (0.005)
Neighbouring region [β_2]	0.229 ^{***} (0.021)	0.710 ^{***} (0.017)
Country border effects [β_3]	-0.062 ^{***} (0.016)	-1.058 ^{***} (0.016)
Language area effects [β_4]	-0.163 ^{***} (0.013)	-0.740 ^{***} (0.014)
Technological Distance [β_5]	-0.300 ^{***} (0.018)	-1.535 ^{***} (0.023)
Core region [β_6]	0.049 ^{***} (0.022)	0.010 (0.026)
Number of significant time effects	7	5
Number of origin spatial filters	32	39
Number of destination spatial filters	29	47
Constant [α_0]	-6.465 ^{***} (0.041)	-2.426 ^{***} (0.041)
Dispersion parameter	19.838 ^{***} (0.253)	2.723 ^{***} (0.045)
LR Test (spatial filters)	1,323.59 ^{***}	4,913.28 ^{***}
LR Test (random effects)	19,0241.7 ^{***}	42,538.3 ^{***}
LR Test (overdispersion)	454,847.2 ^{***}	764,341.7 ^{***}
Log Likelihood	-879,636.9	-415,169.8

Notes: ***significant at the 0.001 significance level

The LR Test (spatial filter) is a Likelihood Ratio test that compares the model fit of the spatially filtered model against the unfiltered model versions. The test statistic is significant for both models. Thus the spatially filtered model specification is appropriate. The LR Test (random effects) is a Likelihood Ratio test that compares the panel estimator with the pooled estimator. The significant values confirm the importance of a random effect specification. The LR Test (overdispersion) is the Likelihood Ratio Test that compares the random effects negative binomial model to the random effects Poisson specification (see Long and Freese 2001). A significant value points to the existence of overdispersion, namely, the negative binomial specification is to be preferred to the Poisson specification.

Technological distance (β_5) is the most important determinant for cross-region R&D collaborations in both networks, and, by this, earlier results by Scherngell and Barber (2009 and 2011) or Fischer et al. (2006) are confirmed. However, the effect is much stronger in the

co-patent network, which is to be expected since co-patent networks are more application oriented, where specific technologies and technological devices are more important. Furthermore, the FPs are intended to support in particular interdisciplinary knowledge production.

Core region effects (β_6) are significant – though rather small – in the FP network, while even insignificant in the co-patent network, pointing to the fact that in the FP network a geographical core of regions is disproportionately participating in FP collaborations as compared to the average region, i.e. non-core regions have a lower probability to get connected (Scherngell and Lata 2012). This may be related to the fact that organisations located in core regions may be more experienced in collaborations with different public stakeholders, and, thus, in conducting FP projects than organisations in peripheral regions, while such assets may not be necessarily essential for organisations participating co-patent collaborations. However, the effect is rather small as compared to other separation effects.

Overall, in the context of our focus on integration in European R&D, we can infer that integration is much higher in the FP network than in the co-patent, as most of the separation variables exert a higher negative effect. This result has been expected, since more applied oriented, competitive research is subject to a minor group of actors often located within one region. However, having in mind the ERA goal of progress towards more integration in European R&D, covering different phases of R&D, one may conclude that barriers hampering collaborations in the co-patent network – for instance language barriers or country borders – should be addressed more thoroughly. This may be addressed by education programs for overcoming language barriers or policy initiatives that remove institutional hurdles for collaborations in patenting, though, one have to be clear that due to the competitive character of this type of research, such patterns may never fully disappear.

However, in order to be able to gain empirical insight into progress towards more integration, we need to reflect on time trends. For this reason we look at interaction terms between selected separation variables and our time dummies. Table 3 presents the results for these interaction terms in the two networks for the years 2000 to 2005. The most striking result is that all separation variables accounting for spatial effects significantly decline in the FP-network, i.e. the FP network becomes more geographically integrated over the observed time period. This cannot be observed for the co-patent network. In particular for the years 2004

and 2005 we cannot identify a significant interaction effect between time and spatial separation variables, i.e. progress towards more integration cannot be observed, while this progress can be clearly observed for the FP network.

Table 3. Time trends for identifying distinct geographical integration patterns in the networks

Time interaction terms	FP-Network					
	2000	2001	2002	2003	2004	2005
Geographical distance	-0.100*** (0.000)	-0.076*** (0.000)	-0.053*** (0.000)	-0.050*** (0.000)	-0.013*** (0.000)	-0.008*** (0.000)
Neighbouring region	-0.375*** (0.013)	-0.280*** (0.012)	-0.188*** (0.012)	-0.180*** (0.012)	-0.055*** (0.011)	-0.009 (0.011)
Country border effects	-0.507*** (0.003)	-0.382*** (0.002)	-0.264*** (0.002)	-0.249*** (0.002)	-0.065*** (0.001)	-0.040*** (0.001)

Time interaction Terms	Co-Patent Network					
	2000	2001	2002	2003	2004	2005
Geographical distance	-0.016*** (0.002)	-0.012*** (0.002)	-0.008*** (0.002)	-0.005** (0.002)	-0.000 (0.002)	-0.000 (0.002)
Neighbouring region	0.047** (0.025)	0.021 (0.024)	0.073** (0.025)	0.036 (0.024)	0.029 (0.024)	-0.003 (0.024)
Country border effects	-0.131*** (0.019)	-0.093*** (0.019)	-0.068*** (0.019)	-0.061*** (0.091)	0.004 (0.019)	0.001 (0.018)

***significant at the 0.001 significance level, **significant at the 0.01 significance level

7 Conclusions

The focus of this study has been on the nature of integration processes in European R&D. More specifically we have shifted emphasis to the investigation of the geographical dynamics of two different types of R&D collaboration networks across Europe, namely co-patent networks and project based R&D networks within the EU Framework Programmes (FPs). Adopting a spatially filtered panel version of the Negative Binomial spatial interaction model, we have identified and compared geographical, technological, institutional and cultural effects that influence the probability for collaboration activities in the different collaboration networks over time, and, by this, have provided novel evidence on integration processes in European R&D.

The results are quite astonishing, as – given the complex nature of the data we are investigating – the advanced methodological approach employed in this study seems to be a very effective way to disclose geographical dynamics of R&D collaboration patterns in the two networks under consideration. The most elemental and important result, both in the context of the literature on the geography of innovation as well as in a European policy context, is that integration in FP networks seems to be much higher than in the co-patent network. This is underpinned by the strong intra-national character of the co-patent network in contrast to the FP network, as well as the higher geographical localisation of co-patent collaboration activities within narrow geographical boundaries. These results may on the one hand be explained by the different nature of the knowledge creation process in the two networks, but also by policy related circumstances, in that the FP programmes explicitly foster integration processes, and at the same time more policy efforts should be envisaged that ease collaboration in more applied oriented research.

Methodologically, the study is interesting as it breaks new ground by estimating a panel version of the Negative Binomial spatial interaction model accounting for spatial autocorrelation of flows. Though robustness of the model may be tested further, the methodological approach seems to be an important contribution to the debate on spatial autocorrelation issues of flows, applied to a panel data structure posing additional modelling requirements that have been applied in this study.

Some ideas for future research come to mind. *First*, the estimation of time trends, for instance by means of a dynamic version of the spatial interaction model, is a core subject for future research, requiring both theoretical as well as computational advancements. *Second*, the inclusion of other types of R&D networks in the comparative analysis, in particular co-publication networks, is essential to complement the results provided by the current study.

Acknowledgments This work has been funded by the FWF Austrian Science Fund (Project No. P21450).

References

- Baltagi B (2008) *Econometric Analysis of Panel Data*, Fourth Edition, John Wiley, Chichester
- Begg I (2010) Cohesion or confusion: a policy searching for objectives. *Journal of European Integration* 32, 77-96
- Berglund S. and Karlström A. (1999): Identifying local spatial association in flow data, *Journal of Geographical Systems* 1(3), 219-236
- Boyer, R (2009) From the Lisbon Agenda to the Lisbon Treaty: national research systems in the context of European integration and globalization, In Delanghe, H., Muldur U. and Soete L. (eds): *European Science and Technology Policy. Towards Integration or Fragmentation?*, Cheltenham, Edward Elgar, pp. 101-126
- Breschi S and Cusmano L (2004) Unveiling the texture of a European research area: Emergence of oligarchic networks under EU Framework Programmes, *International Journal of Technology Management*, Special Issue on Technology Alliances, 27(8), 747-772
- 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, pp. 160-174
- Bröcker J (1989) Partial equilibrium theory of interregional trade and the gravity model, *Papers in Regional Science* 66, 7-18
- Caloghirou Y, Vonortas N and Ioannide S (2002) Science and technology policies towards research joint ventures, *Science and Public Policy* 29, 82-94
- Brunet, R. (2002): Lignes de force de l'espace Européen, *Mappemonde* 66, 14-19
- Cameron AC and Trivedi PK (1998) *Regression Analysis of Count Data*. Cambridge [UK] and New York [US], Cambridge University Press
- Castells M (1996) *The Rise of the Network Society*. Malden, Blackwell
- CEC (Commission of the European Communities) (2007) Green paper 'The European Research Area: New Perspectives'. {SEC(2007) 412}, COM(2007)161 final, Brussels, 4 April 2007
- CEU (Commission of the European Union) (2008) Council conclusions on the definition of a "2020 vision for the European Research Area". 16012/08 RECH 379 COMPET 502, Brussels, 9 December 2008
- Cohen WM and Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35(1), 128-152
- CORDIS (2006) FP5 - EC Programme: Maximum Amounts and Breakdown (1998-2002), available from <http://cordis.europa.eu/fp5/src/budget.htm>
- Cowan R (2004) Network models of innovation and knowledge diffusion. Unu-MERIT working paper series, 2004-016
- Cowan R., Zinovyeva n. (2009) Papers or patents : Channels of University effect on regional innovation. Working Paper, p.36.
- Chun, Y. (2008) Modeling network autocorrelation within migration flows by eigenvector spatial filtering, *Journal of Geographical Systems* 10, 317-344
- Chun, Y. and Griffith, D. (2011): Modeling network autocorrelation in space-time migration flow data: An eigenvector spatial filtering approach, *Annals of the Association of American Geographers*, 101(3), 523-536
- Delanghe, H., Muldur U. and Soete L. (2009) (eds): *European Science and Technology Policy. Towards Integration or Fragmentation?*, Cheltenham, Edward Elgar
- Ejermo, O and Karlsson, C (2006) Interregional inventor networks as studied by patent coinventorships, *Research Policy*, vol. 35(3), pages 412-430
- European Council (1998) Council decision of 22 December 1998 concerning the rules for the participation of undertakings, research centres and universities and for the dissemination of research results for the implementation of the fifth framework programme of the European Community (1998-2002). Available from <http://cordis.europa.eu/fp5/-management/particip/r-ecdec-part.htm#Section%203%20-%20Contracts>
- European Council (2000) Presidency conclusions, Lisbon European Council, 23-24 March 2000.

- European Commission (2011) Progress Report on the Europe 2020 Strategy, Brussels, European Commission, Com (2011) 815 final
- Fischer MM and Wang J (2011) *Spatial Data Analysis. Models, Methods and Techniques*. Heidelberg, Dordrecht, London and New York, Springer
- Fischer MM, Scherngell T, 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 MM and Griffith D (2008) Modeling spatial autocorrelation in spatial interaction data: An application to patent citation data in the European Union, *Journal of Regional Science* 48, 969–989
- Fischer MM and LeSage (2010) Spatial econometric methods for modeling origin destination flows. In: Fischer Manfred M, Getis Arthur (ed.): *Handbook of Applied Spatial Analysis*, Berlin Heidelberg, Springer, pp. 409-432
- Goetze, C (2010) An empirical enquiry into co-patent networks and their stars: The case of cardiac pacemaker technology, *Technovation* 30, pp.436-446
- Granstand O (1998) Towards a theory of the technology-based Firm. *Research Policy* 27(5): 465-489
- Griffith D (2003) *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization*, Berlin Heidelberg, Springer-Verlag
- Griffith D (2009) Spatial Autocorrelation in Spatial Interaction: Complexity-to-Simplicity in Journey-to-Work Flows. In: A. Reggiani and P. Nijkamp (ed): *Complexity and Spatial Networks: In Search of Simplicity*, Berlin New York, Springer, pp. 221-237
- Griffith D (1996) Spatial autocorrelation and eigenfunctions of the geographic weight matrix accompanying georeferenced data, *The Canadian Geographer* 40, 341-367
- Griffith (2011): Visualizing analytical spatial autocorrelation components latent in spatial interaction data: An eigenvector spatial filter approach, *Computers, Environment and Urban Systems* 35, 140-149
- Griliches, Z, (1984) R & D, Patents, and Productivity, NBER Books, National Bureau of Economic Research.
- Grossman, G.M. and Helpman, E. (1991) Quality leaders in the theory of growth, *Quarterly Journal of Economics* 106, 557-86
- Guzzetti, L (2009) The european research area idea in the history of community policy-making , In Delanghe, H., Muldur U. and Soete L. (eds): *European Science and Technology Policy. Towards Integration or Fragmentation?*, Cheltenham, Edward Elgar, pp. 64-77
- Hoekman, J., Scherngell. T., Frenken, K. and Tijssen, R. (2012): Acquisition of European research funds and its effect on international scientific collaboration, *Journal of Economic Geography* [forthcoming]
- 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
- Hoekman, J., Frenken, K., Tijssen, R.J.W. (2010) Research collaboration at a distance : changing spatial patterns of scientific collaboration in Europe. *Research Policy* 39(5), 662-673.
- Jaffe, A.B., Trajtenberg, M. and Henderson, R. (1993): Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics* 108(3), 577-598
- Katz, JS (1994) Geographical proximity and scientific collaboration, *Scientometrics* 31: 31-43.
- Kogut B (1988): Joint ventures: Theoretical and empirical perspectives, *Strategic Management Journal* 9, 319-332
- LeSage, J and Fischer M.M (2012) Estimates of the Impact of Static and Dynamic Knowledge Spillovers on Regional Factor Productivity, *International Regional Science Review*, 35, pp. 103-127
- LeSage J and Pace K (2008) Spatial econometric modeling of origin-destination flows, *Journal of Regional Science* 48, 941-967
- LeSage J, Fischer MM, Scherngell T (2007) Knowledge spillovers across Europe. Evidence from a Poisson spatial interaction model with spatial effects, *Papers in Regional Science* 86: 393–421

- LeSage, J. and Pace K (2005) Spatial Econometric Modeling of Origin-Destination flows, paper presented at the Regional Science Association International Meeting, Las Vegas
- Lin, G. and Zhang, T.(2007): Loglinear residual tests of Moran's I autocorrelation and their applications to Kentucky breast cancer data, *Geographical Analysis* 39, 293-310
- Long JS, Freese J (2001) Regression models for categorical dependent variables using stata. Stata Corporation, College Station
- Lucas, R.E. (1988) On the mechanics of economic development, *Journal of Monetary Economics* 22, pp.3-42
- Luukkonen,T (2010) The European Research Council in the European research Area, EURECIA D7.1 ,
- Lundvall BA (1992) *National innovation systems*. Pinter, London
- Maggioni MA and Uberti TE (2007) Inter-regional knowledge flows in Europe: An econometric analysis. In Frenken K (ed.): *Applied Evolutionary Economics and Economic Geography*, Cheltenham, Edward Elgar, pp.230-255
- Maggioni MA, Nosvelli M, Uberti TA (2007) Space versus networks in the geography of innovation: A European analysis. *Papers in Regional Science* 86: 471-493
- Maggioni MA and Uberti TE (2009) Knowledge networks across Europe: which distance matters?,*The Annals of Regional Science* 43: 691-720
- Makri, M, Hitt, M, and Lane, P (2010) Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions, *Strategic Management Journal* 31, pp. 602-628
- Moreno R, Paci R and Usai S (2005) Spatial spillovers and innovation activity in European regions, *Environment and Planning A* 37, 1793-1812
- Owen-Smith, J. and Powell, W. W. (2002): Knowledge Networks in the Boston Biotechnology Community. Paper presented at the Conference on 'Science as an Institution and the Institutions of Science' in Siena
- Parent O and LeSage J (2012) Determinants of knowledge production and their effects on regional economic growth, *Journal of Regional Science* 52(2), 256-284
- 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(1), 89-104
- Patuelli R, Griffith D, Tiefelsdorf M and Nijkamp P (2011) Spatial Filtering and Eigenvector Stability:Space-Time Models for German Unemployment Data, *International Regional Science Review* 34(2), 253-280
- Pavitt K (2005) Innovation processes. In: Fagerberg J, Mowery DC, Nelson RR (eds) *The Oxford Handbook of Innovation*. Oxford University Press, Oxford, 86-114
- Penrose ET (1959) *The Theory of the Growth of the Firm*. Oxford, Oxford University Press
- Peterson J and Sharp M (1998) Technology policy in the European Union. Macmillan Press, London
- Polt, W., Vonortas N. and Fisher, R. (2008), Innovation Impact, Final report to the European Commission, Brussels: DG Research.
- Ponds R, Van Oort F and Frenken K (2007) The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86: 423-443
- Powell WW and Grodal S (2005) Networks of innovators, In: Fagerberg J, Mowery DC and Nelson RR (eds.), *The Oxford Handbook of Innovation*, Oxford, Oxford University Press, 56-85
- Roediger-Schluga T, Barber M (2006) The structure of R&D collaboration networks in the European Framework Programmes. Unu-MERIT Working Paper Series 2006-36, Economic and Social Research and Training Centre on Innovation and Technology, United Nations University, Maastricht
- Roediger-Schluga T, Barber M (2008) R&D collaboration networks in the European Framework Programmes: Data processing, network construction and selected results. *International Journal of Foresight and Innovation Policy* 4: 321-347
- Romer PM (1990) Endogenous technological change. *Journal of Political Economy* 98: 71-102
- 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*, [forthcoming]

- Scherngell T and Barber MJ (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 MJ (2009) Spatial interaction modelling of cross-region R&D collaborations. Empirical evidence from the 5th EU Framework Programme, *Papers in Regional Science* 88, 531-645
- Scherngell T and Hu Y (2011) Collaborative knowledge production in China. Regional Evidence from a Gravity Model Approach, *Regional Studies* 45, 755-772
- Sen, A. and Smith, T.E. (1995): Gravity Models of Spatial Interaction Behaviour. Heidelberg, Berlin and New York, Springer
- Svanfeldt, C. (2009) A European Research Area built by the Member States? In Delanghe, H. Muldur, U. and Soete, L. (eds.), *European science and technology policy: towards integration or fragmentation?*, Edward Elgar Publishing, 44-63
- Tel Wal, A.L.J. and Boschma, R. A. (2008) Applying social network analysis in economic geography: framing some key analytic issues, *The Annals of Regional Science* 43, 739-756
- Tiefelsdorf, M. and Boots, B.N. (1995) The exact distribution of Moran's I, *Environment & Planning A* 27, 985-999
- Williamson OE (1975) *Markets and Hierarchies*. New York, Free Press

Appendix A

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 255 NUTS-2 regions located in the EU-25 member states (except Cyprus and Malta) plus Norway and Switzerland. 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:

<i>Austria:</i>	Burgenland, Kärnten, Niederösterreich, Oberösterreich, Salzburg, Steiermark, Tirol, Vorarlberg, Wien
<i>Belgium:</i>	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
<i>Czech Republic:</i>	Jihovýchod, Jihozápad, Moravskoslezsko, Praha, Severovýchod, Severozápad, Střední Morava, Střední Čechy
<i>Denmark:</i>	Danmark
<i>Estonia:</i>	Eesti
<i>Finland:</i>	Åland, Etelä-Suomi, Itä-Suomi, Länsi-Suomi, Pohjois-Suomi
<i>France:</i>	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
<i>Germany:</i>	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
<i>Greece:</i>	Anatoliki Makedonia, Thraki; Attiki; Ipeiros; Voreio Aigaiο; Dytiki Ellada; Dytiki Makedonia; Thessalia; Ionia Nisia; Kentriki Makedonia; Kriti; Notio Aigaiο; Peloponnisos; Sterea Ellada
<i>Hungary:</i>	Dél-Alföld, Dél-Dunántúl, Észak-Alföld, Észak-Magyarország, Közép-Dunántúl, Közép-Magyarország, Nyugat-Dunántúl
<i>Ireland:</i>	Border, Midland and Western; Southern and Eastern

<i>Italy:</i>	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
<i>Latvia:</i>	Latvija
<i>Lithuania:</i>	Lietuva
<i>Luxembourg:</i>	Luxembourg (Grand-Duché)
<i>Netherlands:</i>	Drenthe, Flevoland, Friesland, Gelderland, Groningen, Limburg (NL), Noord-Brabant, Noord-Holland, Overijssel, Utrecht, Zeeland, Zuid-Holland <i>Norway:</i> Agder og Rogaland, Hedmark og Oppland, Nord-Norge, Oslo og Akershus, Sør-Østlandet, Trøndelag, Vestlandet
<i>Poland:</i>	Dolnośląskie, Kujawsko-Pomorskie, Lubelskie, Lubuskie, Łódzkie, Mazowieckie, Małopolskie, Opolskie, Podkarpackie, Podlaskie, Pomorskie, Śląskie, Świętokrzyskie, Warmińsko-Mazurskie, Wielkopolskie, Zachodniopomorskie
<i>Portugal:</i>	Alentejo, Algarve, Centro (P), Lisboa, Norte
<i>Slovakia:</i>	Bratislavský kraj, Stredné Slovensko, Východné Slovensko, Západné Slovensko
<i>Slovenia:</i>	Slovenija
<i>Spain:</i>	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
<i>Sweden:</i>	Mellersta Norrland, Norra Mellansverige, Småland med öarna, Stockholm, Sydsverige, Västsverige, Östra Mellansverige, Övre Norrland
<i>Switzerland:</i>	Espace Mittelland, Nordwestschweiz, Ostschweiz, Région lémanique, Ticino, Zentralschweiz, Zürich
<i>United Kingdom:</i>	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 & Warwickshire; 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