

---

# “Research Networks and Inventors’ Mobility as Drivers of Innovation: Evidence from Europe”

Ernest Miguélez & Rosina Moreno  
AQR-IREA Universitat de Barcelona

---



Institut de Recerca en Economia Aplicada Regional i Pública  
*Research Institute of Applied Economics*

**Universitat de Barcelona**

Av. Diagonal, 690 • 08034 Barcelona

WEBSITE: [www.ub.edu/irea/](http://www.ub.edu/irea/) • CONTACT: [irea@ub.edu](mailto:irea@ub.edu)

The Research Institute of Applied Economics (IREA) in Barcelona was founded in 2005, as a research institute in applied economics. Three consolidated research groups make up the institute: AQR, RISK and GiM, and a large number of members are involved in the Institute. IREA focuses on four priority lines of investigation: (i) the quantitative study of regional and urban economic activity and analysis of regional and local economic policies, (ii) study of public economic activity in markets, particularly in the fields of empirical evaluation of privatization, the regulation and competition in the markets of public services using state of industrial economy, (iii) risk analysis in finance and insurance, and (iv) the development of micro and macro econometrics applied for the analysis of economic activity, particularly for quantitative evaluation of public policies.

IREA Working Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. For that reason, IREA Working Papers may not be reproduced or distributed without the written consent of the author. A revised version may be available directly from the author.

Any opinions expressed here are those of the author(s) and not those of IREA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

---

### *Abstract*

---

We investigate the importance of the labour mobility of inventors, as well as the scale, extent and density of their collaborative research networks, for regional innovation outcomes. To do so, we apply a knowledge production function framework at the regional level and include inventors' networks and their labour mobility as regressors. Our empirical approach takes full account of spatial interactions by estimating a spatial lag model together, where necessary, with a spatial error model. In addition, standard errors are calculated using spatial heteroskedasticity and autocorrelation consistent estimators to ensure their robustness in the presence of spatial error autocorrelation and heteroskedasticity of unknown form. Our results point to the existence of a robust positive correlation between intra-regional labour mobility and regional innovation, whilst the relationship with networks is less clear. However, networking across regions positively correlates with a region's innovation intensity.

*JEL classification:* C8, J61, O31, O33, R1

*Keywords:* inventors' mobility, networks of co-inventors, knowledge production function, spatial econometrics, European regions

Ernest Miguelez is Predoctoral Scholar (FPU), Ministry of Education, at University of Barcelona. Department of Econometrics, Statistics and Spanish Economy. E-mail: emiguelez@ub.edu.  
Rosina Moreno is Full Professor at University of Barcelona Department of Econometrics, Statistics and Spanish Economy. E-mail: rmoreno@ub.edu.

### *Acknowledgements:*

---

Part of this work was carried out while Ernest Miguélez was visiting the Kiel Institute for the World Economy (Kiel, Germany). The use of the Institute's facilities is gratefully acknowledged. The authors would also like to share their appreciation of helpful comments received from Wan-Hsin Liu, Aida Solé-Auró, and the participants of the "XXXIV Simposio de la Asociación Española de Economía" (Valencia, December 10<sup>th</sup>-12<sup>th</sup> 2009). The work made by Ismael Gómez-Miguélez is sincerely appreciated. We also acknowledge financial support from the Ministerio de Ciencia e Innovación, ECO2008-05314 and Ernest Miguélez, from the Ministerio de Educación, AP2007-00792. However, any mistake or omission remains ours.

---

## 1. Introduction

Knowledge diffusion between individuals and firms is critical for innovation and growth (GROSSMAN and HELPMAN, 1991; LUCAS, 1988; ROMER, 1986, 1990). At the firm and geographical levels knowledge is known to flow through a variety of mechanisms (DÖRING and SCHNELLENBACH, 2006), among which the labour mobility of highly skilled personnel and networks of research collaboration are considered pivotal.

The goal of this paper is to analyse the contribution made by these collaborative networks and the labour and geographical mobility of inventors to the process of knowledge creation. To do so, we investigate the quantitative relationship between these two features of the local labour market and regional innovative capability. Specifically, our empirical analysis is built upon a regional knowledge production function (KPF hereafter) framework, which is used to model the aforementioned relationship for a large sample of European NUTS-2 regions.

By drawing on patent data and computerized algorithms to identify individual inventors, we construct a large dataset of individuals containing information regarding their personal address(es), their patenting history, the owners of their patents (be it a firm, a university or other public institution, or the inventors themselves), and the co-authors of their patents –among other details. To the best of our knowledge, very few studies have examined the influence of these features on regional innovation because of a lack of data on single, individual inventors.

Among the questions addressed in this study are the following: What is the contribution of inventor networking and inventors' labour mobility to the regional intensity of patenting? What do these market-based flows of knowledge contribute, above and beyond the contribution made by non-market knowledge externalities, to the variation in regional innovative capability? Do cross-regional mobility and cross-regional networking play an important role? What impact is attributable to mobility and networking once spatial interactions have been controlled for?

Our choice of the term knowledge “flows” instead of, for instance, knowledge “spillovers” or “externalities” is not accidental and requires an explanation before we proceed. As BRESCHI and LISSONI (2009) point out, knowledge may flow locally, but not necessarily as an externality. When knowledge is transferred following the hiring of an employee from another firm, or by means of a cooperation agreement, market mechanisms are operating and no externalities or spillovers can be said to exist. Clearly, these flows may include pure externalities if all the knowledge transferred is not compensated accordingly, but in all instances it will only constitute a marginal part. For this reason,

in our empirical approach we seek to disentangle the effect of pure non-market externalities from that of market-based knowledge flows.

Our econometric approach takes full account of the spatial structure of our data by applying spatial econometric techniques. Unlike previous studies of innovation, however, we do not use a spatial lag model to the exclusion of the error model or vice-versa, but rather we allow for a rich spatial specification. To achieve this, we first estimate the spatial lag model and then we perform the non-parametric heteroskedasticity and autocorrelation consistent (HAC) estimator of the variance-covariance (V-C) matrix in a spatial framework (following KELEJIAN and PRUCHA, 2007). To the best of our knowledge, no previous paper has dealt with the estimation of a knowledge production function using both a spatial lag model and the robust spatial HAC estimation of the V-C matrix.

The rest of the paper is organized as follows: section 2 reviews the literature on inventor networking and mobility, and their relationship with knowledge diffusion and creation; section 3 describes the model and the hypotheses we propose; section 4 presents the data; whilst section 5 includes preliminary OLS results. Section 6 extends our analysis towards spatial econometrics and cross-regional mobility and networking, and finally, section 7 presents our conclusions and identifies certain limitations in our approach.

## 2. Background

The literature on collaborative research networks, and their impact on knowledge diffusion and innovation, has expanded greatly in recent years.<sup>1</sup> This is particularly true in the case of networks of co-inventors thanks to the availability of relevant data (co-patent data). Part of this literature has been devoted to explaining the determinants of these collaborative patterns (HOEKMAN et al. 2009; MAGGIONI and UBERTI, 2008), while a further important line has focused on networks as mechanisms for inter-regional R&D spillovers (KROLL, 2009; PONDS et al., 2007, 2009), and, in particular, networks as the means by which knowledge diffuses between individuals and across firms (BRESCHI and LISSONI, 2004, 2006; 2009; GOMES-CASSERES et al., 2006; SINGH, 2005).

SINGH (2005) finds strong evidence in the US that the existence of interpersonal ties in the form of co-patents increases the probability of knowledge flows, as measured by patent citations. Indeed, he

---

<sup>1</sup> Recent special issues on the subject include: “Spatial knowledge diffusion through collaborative networks” Guest editors: CORINNE AUTANT-BERNARD, JACQUES MAIRESSE and NADINE MASSARD, *Papers in Regional Science* 2007, 86(3): 341-525; and, more specifically on the subject of networks of co-inventors, the special issue, “Embedding network analysis in spatial studies of innovation”. Guest editor: EDWARD M. BERGMAN, *The Annals of Regional Science*, 2009, 43(3): 559-833.

hints at the fact that there is evidence that such ties are important in determining patterns of intra-regional and intra-firm knowledge flows. SINGH claims that geography matters especially because interpersonal networks tend to be regional in nature (Op. cit.). Similar results are found by BRESCHI and LISSONI (2004) for Italy. In the same line, recent findings by BRESCHI and LISSONI (2009), using patent applications to the European Patent Office (EPO) from certain technological fields (biotechnology, organic chemistry, and pharmaceuticals), similarly suggest that networking activity across firms (and locations) is in large part responsible for the localisation of knowledge flows, indicating that the residual effect of non-market externalities is not as great as was previously believed.

All these studies stress the importance of networks as knowledge transmission, and hence creation, mechanisms. Co-location and shared space are reported as being neither necessary nor sufficient for knowledge flows, but rather it is social distance, or the quality and quantity of interactions between those individuals most closely involved with innovative activities, which appears to be critical for the effective diffusion of knowledge (BOSCHMA, 2005). If this were to be the case, the features of the inventors' network structure at any given location should play a significant role in regional innovation outcomes. In this sense, a number of empirical analyses have recently been conducted in a knowledge production function framework by BETTENCOURT et al. (2007a,b), FLEMING et al. (2007), and LOBO and STRUMSKY (2008) for the case of US metropolitan statistical areas (MSAs). Broadly speaking, various features of the inventors' network are found to have different and significant effects on creativity. While the scale and extent of networks are reported to have a positive impact, the density of these networks seems to have a negative effect, which has been attributed to the existence of superfluous information (BETTENCOURT et al., 2007a,b; LOBO and STRUMSKY, 2008). However, the importance of inventors' networks has not been found to be equally significant in all the analyses conducted (FLEMING et al., 2007), and further research in this field is still required. But, yet, empirical studies of inventors' networks and the evidence of their benefits remain scarce.

Our research is closely related to these recent studies as we also employ a KPF to assess the effect of networks and their particular structure on aggregate innovation outcomes. However, here we apply the production function to European regions as opposed to US MSAs.<sup>2</sup> Moreover, unlike in these earlier studies, here spatial relationships across regions are taken into full account by applying spatial econometric techniques. Yet, no significantly different results are expected a priori with those previously reported for the US.

---

<sup>2</sup> Only very recently a working paper by KROLL (2009) examined the effect of the characteristics of inventors' networks on innovation outcomes; however, both the features it focuses on and the approach it adopts are slightly different from ours.

Similarly, earlier studies have examined how the labour mobility of inventors acts as a key mechanism in the diffusion of knowledge (ALMEIDA and KOGUT, 1999; ROSENKOPF and ALMEIDA, 2003; SAXENIAN, 1994). This line of research however has been less prolific.<sup>3</sup> One strand of this literature has shown the relationship between mobility and the flow of knowledge as measured by patent citations, as well as the knowledge gain by a firm hiring an inventor from another firm. For example, in a pioneering study, ALMEIDA and KOGUT (1999) show that inter-firm mobility of patent holders in the semiconductor industry of the US influences the local transfer of knowledge across firms. Similar findings are reported in the aforementioned study conducted by BRESCHI and LISSONI (2009) for US inventors in selected technological fields making patent applications to the EPO. In a similar vein, AGRAWAL et al. (2006) stress the idea that once inventors leave their workplace, they will maintain interpersonal ties with their former colleagues which can translate into a citation of their work by these co-workers. In addition, several studies (CRESPI et al., 2007; CORREDORIA and ROSENKOPF, 2006; KIM et al. 2006; SINGH and AGRAWAL, 2009; SONG et al., 2003) have stressed the role of mobility insofar as it increases the hiring firm's use of a hired inventor's prior knowledge.

Parallel to these studies, another line of research has studied mobility by focusing its attention on inventors' performance itself. For instance, mobility-productivity relationships have been studied by HOISL (2007, 2009) for German, by LENZI (2009) for Italian, and by SHALEM and TRAJTENBERG (2008) for Israeli inventors. Broadly speaking, it has been shown that mobility may in fact enhance productivity (HOISL, 2007), although results in that direction are not as robust as one would expect (SHALEM and TRAJTENBERG, 2008), which has been attributed to what have been identified as the short-term costs of mobility.

If an individual's innovative output on moving increases and the new host firms acquire more knowledge and are more efficient in their innovative activities, the innovative capacity of a region as a whole should increase as the degree of inventors' mobility rises across firms within a region. To the best of our knowledge, there have been few empirical attempts at quantifying this feature of the local labour market for inventors -in other words, the degree of job-to-job mobility of inventors - on regional innovation outcomes, and as such this constitutes one of the main contributions of this paper.

---

<sup>3</sup> The number of such studies has been limited by data constraints, and only recently have researchers begun using patent data to identify inventors and their subsequent movements. The reason for this limitation is basically the cumbersome nature of singling out inventors on the basis of the names disclosed in patent documents.



### 3. Model and hypotheses

In order to meet the goals identified in the introduction to this article, we apply the knowledge production function framework at the regional level. First used in the seminal studies of GRILICHES (1979) and HAUSMAN et al. (1984) at the firm level, this framework was subsequently extended by JAFFE (1986, 1989) to the regional level. Since then, the regional KPF has been the usual tool for assessing the influence of regional innovation efforts on regional innovation output. However, many empirical exercises have sought to test this relationship by simultaneously considering a number of regional structural characteristics, which would also tend to have a positive impact on regional innovation output. All in all, the KPF has been widely estimated for the case of the US (ACS et al., 2002; ANSELIN et al., 1997) and Europe (ACKÇOMACK and TER WEEL, 2009; BOTTAZZI and PERI, 2003; DE DOMINICIS et al, 2007; MORENO et al., 2005a).

Our point of departure is the simplest specification of this model:

$$Y = A \cdot f(K, L), \quad (1)$$

where  $Y$  is the innovative output of a given region, which depends on two knowledge inputs, i.e., the knowledge capital of a given region ( $K$ ), and the R&D employment of that region ( $L$ ) -see GRILICHES (1979), footnote 3. Moreover, we allow the innovation output to depend on an R&D productivity index<sup>4</sup> for that region,  $A$ . For simplicity, it is assumed that the KPF follows a Cobb-Douglas functional form and that knowledge inputs (knowledge capital and R&D employment) show constant returns to scale:

$$Y_i = C \cdot A_i^\delta \cdot K_i^\beta \cdot L_i^\alpha, \quad 0 < \alpha < 1, \quad 0 < \beta < 1, \quad (2)$$

where  $i$  denotes the region, and  $C$  the constant term capturing the impact of all common factors affecting innovation. In order to guarantee reliable results in our estimations, the regional observations need to be comparable in size, so we pre-multiply equation 2 by a factor of  $1/N$ , where  $N$  is the total population of each region.

$$y_i = C \cdot A_i^\delta \cdot k_i^\beta \cdot l_i^\alpha, \quad (3)$$

---

<sup>4</sup> The R&D productivity index in the knowledge production function would be equivalent to the Total Factor Productivity index in a production function.



where  $y = Y/N$ ,  $k = K/N$ , and  $l = L/N$ . In this way, the innovative intensity of regions (patent applications per capita) depends on the stock of knowledge per capita and the number of R&D employees as a proportion of the total population.

From a neoclassical perspective, all kind of R&D efforts will systematically lead to a larger number of inventions, so regions investing more heavily in R&D will tend to innovate more. However, this argument overlooks the importance of a set of factors that actually account for how innovation is generated at the regional level (RODRÍGUEZ-POSE and CRESCENZI, 2008). To capture a variety of returns that might affect innovation outcomes, the R&D productivity index is assumed to be a function of a number of (Hicks neutral) control variables. In addition to pure externalities and other controls, we assume this index to be affected by the level of inventors' labour mobility within a given region, and the scale and density of its collaborative research networks.

Below, we set out the hypotheses that we seek to test herein:

*H1. Collaborative research networks of inventors boost regional innovation capability since such networks facilitate the diffusion of knowledge between individuals.*

The rationale behind this is that the simple cross-fertilization of previously unconnected ideas will lead to better knowledge outputs (HOEKMAN et al., 2009) and that individuals connected within a collaborative framework are more willing to learn from each other than is the case of isolated inventors. Thus, close network links should prove more useful in transferring complex knowledge (COWAND and JONARD, 2004), especially that with a high component of "tacitness" (SINGH, 2005). Additionally, professional relationships of this nature enhance trust and cooperative behaviour between individuals -and hence raising the level of social capital - which has been shown to be a further element in innovation and knowledge transmission (AKÇOMAK and TER WEEL, 2009; DE CLERCQ and DAKHLI, 2004; MIGUÉLEZ et al., 2008).

In line with the literature (see Section 2), we employ social network analysis (SNA) tools to investigate empirically the quantitative relationship between inventors' collaborations and levels of inventiveness.<sup>5</sup> We are particularly interested in measuring two aspects of inventors' networks: first, the scale of these networks, i.e., whether a greater number of social ties, or just simply more inventors involved in co-innovations, are beneficial for inventive intensity- what we wish to know is

---

<sup>5</sup> Aside from the aforementioned studies, SNA has been widely applied to collaboration in research and innovation studies, although a review of detailed methodological contributions falls outside the scope of this paper. In fact, in recent years many contributions have been made to economics and economic geography using SNA tools, most notably BALCONI et al. (2004), BRESCHI and CATALINI (2009), and TER WAL and BOSCHMA (2009). For a more complete theoretical discussion of the methods and applications of SNA, see WASSERMAN and FAUST (1994).

whether a greater number of inventors involved in collaborative research will result in higher innovation outcomes per capita. Here, we use three distinct variables to proxy the scale of the networks in a community – however, we are reluctant to include the three variables at the same time in our regressions because of their high degree of collinearity and, thus, three separate models are estimated. Clearly, a positive effect on creativity is expected.

The second aspect we are concerned with is the strength of the inventors' community ties, measured as the network density. The naïve, expected effect of density on innovation is positive. However, we should bear in mind GRANOVETTER's (1985) warning that overly strong interpersonal ties might well hamper innovation because of the fact that, at some point, the information flowing across those ties becomes redundant. As we show in Section 5 below, the findings of our empirical analysis of network density are in line with those encountered for the case of the US.

*H2. The mobility of inventors within the local labour market of a region enhances innovative intensity since their mobility acts as a source of knowledge flows and human capital externalities.*

Individuals are by definition the main carriers of knowledge. Their movement across firms must therefore contribute to knowledge exchange between firms (BOSCHMA et al., 2009). Skilled workers take their knowledge with them and share it in a new workplace with their new colleagues, at the same time as they provide their new employer with this knowledge. In return, they acquire new knowledge from their new colleagues, establish new links and social networks for future collaborations based on trust and, in general, promote new combinations of knowledge (TRIPPL and MAIER 2007).

Given these two hypotheses, the R&D productivity index is a function of the following variables

$$A_i = g(A_C, MOB_i, NET_i, DENS_i, Z_i, X_i), \quad (4)$$

where *MOB* is our measure of mobility, *NET* is the scale of the networks in a given region, *DENS* is a measure of the density of regional networking, *Z* is a set of variables controlling for the existence of pure externalities derived from agglomeration economies, specialization, urbanization, and market access; and *X* is a set of technical controls to capture differences in economic structures across regions. Moreover,  $A_C$  is a country-specific effect capturing both institutional differences across countries affecting innovation and the existence of so-called national systems of innovation. Thus, equation 3 becomes

$$y_i = C \cdot [g(A_C, MOB_i, NET_i, DENS_i, Z_i, X_i)]^\delta \cdot k_i^\beta \cdot h_i^\alpha. \quad (5)$$

We model the R&D productivity index as a simple multiplicative function of location-specific features:

$$A_i = A_C \cdot MOB_i^{\phi_1} \cdot NET_i^{\phi_2} \cdot DENS_i^{\phi_3} \cdot Z_i^{\phi_j} \cdot X_i^{\phi_n}. \quad (6)$$

Expressing this in log form yields the following specification,

$$\ln A_i = \ln A_C + \phi_1 \cdot \ln MOB_i + \phi_2 \cdot \ln NET_i + \phi_3 \cdot \ln DENS_i + \phi_j \cdot \ln Z_i + \phi_n \cdot \ln X_i. \quad (7)$$

Expressing 3 in log form and substituting 7 into the resulting expression, gives the following empirical model, to which we now add an (a priori) random error term,  $\varepsilon_i$ :

$$\begin{aligned} \ln y_{it} = & \xi + \beta \cdot \ln k_{it-1} + \alpha \cdot \ln l_{it-1} + \omega_1 \cdot \ln MOB_{it-1} + \omega_2 \cdot \ln NET_{it-1} + \\ & + \omega_3 \cdot \ln DENS_{it-1} + \omega_j \cdot \ln Z_{it-1} + \omega_n \cdot \ln X_{it-1} + \\ & + \omega_N \cdot \ln DUMMY\_COUNTRY\_EFFECT + \varepsilon_i \end{aligned}, \quad (8)$$

where  $\xi = \ln C$ ,  $\omega_r = \phi_r \cdot \delta$ ,  $y_{it} = PATpc_{it}$ ,  $k_{it-1} = STOCK\_RDpc_{it-1}$  and  $l_{it-1} = HRST_{it-1}$ . We also add the subscript  $t$  denoting the time dimension, because although we are estimating a cross-section regression, we should be aware that all the explanatory variables pre-date the period of analysis of the dependent variable and could be assumed, in principle, to be exogenous.

#### 4. Data

In this section we outline how we chose to construct our variables, starting from the dependent and focus variables, and the remaining explanatory regressors.

Innovation intensity is measured by patent applications per capita (PATpc), a variable widely used in the literature to proxy innovation outcomes. As well known, this proxy presents serious caveats since not all inventions are patented, nor do they all have the same economic impact, as they are not all commercially exploitable (GRILICHES, 1991). In spite of these shortcomings, patent data have proved useful for proxying inventiveness as they present minimal standards of novelty, originality and potential profits, and as such are a good proxy for economically profitable ideas (BOTTAZZI and

PERI, 2003). We obtain our patent data from the REGPAT database (January 2009 edition), which is the result of a regionalisation of the PATSTAT dataset (all EPO patents) carried out by the OECD. Since these data are prone to exhibit lumpiness from year to year, we have averaged out patent figures from 2001 to 2003, thereby mitigating the effects of annual fluctuations in this variable, especially in those less populated areas. The same procedure is repeated for the explanatory variables -for the period 1996-2000.<sup>6</sup>

R&D expenditure data are drawn from Eurostat and are used to construct the R&D stock proxy (STOCK\_R&Dpc). The accumulation of R&D is described as  $\Delta R \& D = R \& D_t - \delta R \& D_{t-1}$  -see PERI (2005)- where  $\delta$  is an arbitrary depreciation rate set at 5%. Other depreciation rates, however, do not affect the results. The initial stock of R&D has been calculated using the perpetual inventory method,  $R \& D_0 = R \& D_{t-n} / (g + \delta)$ , where  $R \& D_{t-n}$  is the first year for which the data were available, and  $g$  is the geometric average annual growth rate of R&D expenditure between 1996 and 2000. R&D employment has been proxied using Human Resources in Science and Technology (HRST) data extracted from Eurostat. Ideally, direct data on R&D employment, also available from Eurostat, should be used but the number of missing figures is considerable and so we opted not to use this dataset.

The data for constructing the mobility and network variables are also taken from the REGPAT database. In spite of the vast amount of information contained in patent documents, a single ID for each inventor and anyone else involved is missing. However, in order to draw the mobility and networking history of inventors, we need to identify them individually by name and surname, as well as via the other useful details contained in the patent document. The method chosen for identifying the inventors is therefore of the utmost importance in studies of this nature. Thus, here, we have followed the methodology proposed by MIGUÉLEZ and MIGUÉLEZ (2010), who, in line with a growing number of researchers in the field, suggest several algorithms for singling out individual inventors using patent documents. In the present study, this procedure has been used for a subsample of inventors whose patent applications have been made from one of 17 countries -see Appendix 1- between 1993 and 2002. Additionally, and unlike the dependent variable, we only apply the algorithms to EPO patents filed under the Patent Cooperation Treaty (PCT) as well. This condition was adopted for two main reasons: first, because patents filed during an international phase (the PCT) can be assumed to be more technologically and economically exploitable, since applying for EPO patents under PCT procedures is more expensive and time-consuming. Therefore, the

---

<sup>6</sup> In the case of the mobility and networking variables, however, we do not calculate averages but rather total labour movements and co-patents between 1993 and 2002. We do so because of the very small number of these variables, and indeed we had to extend the period of analysis to ensure sufficient variation. There is no a priori reason to believe that this difference in procedure might affect the spatial differences in our sample and, therefore, our estimation results.

knowledge embedded in these patents is presumably of greater worth for innovative activities. The second reason is related to time constraints, because of the fact that the chosen methodology still requires a considerable amount of manual work to ensure minimum levels of reliability. Eventually, however, the procedure provided us with a sample of 646,678 records, from which we identified 285,831 individual inventors.

Once each inventor has been assigned an individual identification, mobility and network data can be calculated for each region. A “mobile” inventor is broadly defined as an individual that moves across different organisations offering his/her services (BRESCHI and LISSONI, 2009). Therefore, mobility can refer either to labour mobility understood in its strictest sense (an employee leaving a firm to take up a position in a new one), or to that demonstrated by consultants, freelance workers, university inventors, and the like. We assume both to constitute sources of knowledge flows to the extent that in the two instances knowledge is transferred from their former employers or customers to new ones. Mobility is then proxied as the average number of assignees<sup>7</sup> for each inventor identified per region. To make this calculation, we draw on the recent study by THOMA et al. (2009a,b) and the OECD Harmonised Applicants’ Names (HAN) database (October, 2009), aimed at identifying single applicants in patent documents. Additionally, we have to discard all patents with more than one applicant, or assignee, as we were unable to discriminate labour relationships between applicants and inventors in a multi-applicant patent. However, from the initial sample of 646,678, only 59,687 records were discarded.

The design of the network variables is built upon the theory of SNA. Thus, the inventors form the nodes in the network, and these are grouped via edges or ties (in this instance, co-patents) into different components. A component, therefore, is formed by all the inventors in the network that are linked in such a way that a continuous path via the edges can be traced across them. In this way, even if a given pair of inventors has not co-patented an invention, but they share a co-invention with a third inventor, they will form part of the component, i.e., both direct and indirect links make up a component. The largest component is the largest set of inventors that can be linked through direct and indirect collaboration ties.

The scale of the network structure in each region is proxied using three different variables. Connectivity (CONNECTIVITY) measures the absolute number of edges within a community of inventors (a given region). The size of the largest component (SIZE\_LC) is the proportion of inventors in a given region that belong to the largest component, again, of that region. Finally, a

---

<sup>7</sup> In our study, we use the term firm, applicant, or assignee interchangeably to refer to the owner of the patent listed in the patent document - although clearly they are not always the same.

relative measure of the number of ties (edges) to the number of inventors (*TIES\_INV*) is also considered. A positive and significant coefficient is expected both for the mobility and networking variables.

The strength of these ties is proxied by the network density, which is the number of ties between inventors within the region divided by the possible number of ties within that region. Formally,

$$DENS_i = \frac{T}{Q(Q-1)/2}, \quad (9)$$

where  $T$  stands for the number of ties, and  $Q$  is the total number of inventors within a given region. As stressed earlier, the expected effect (be it positive or negative) of innovation density is not so clear a priori.

In our specification, up to four variables are included to capture the effect of pure non-marketable externalities ( $Z$ ) as opposed to market-based knowledge flows (such as mobility and collaborative networks). Thus, gross value added per capita (capturing market access, general economic conditions, and wealth - *GVApc*), the number of technological firms to the total number of manufacturing establishments in that region (capturing agglomeration of firms engaged in patenting activity - *TECH\_FIRMS*), population density (capturing agglomeration economies and urbanization economies -inter-industry economies- *POP\_DENS*), and a technological specialization index (capturing specialization economies -intra-industry economies- constructed using the eight technological classes of the *IP*<sup>8</sup> classification<sup>9</sup> - *SPIN*) are considered in our regressions.

Likewise, a set of four controls ( $X$ ) capture differences in the technological and economic structures across regions: the shares of manufacturing employment in each region (*MAN*), and the shares of biotechnology (*BIOTECH*), organic chemistry (*CHEM*), and pharmaceuticals (*PHARMA*) in their patenting activity, according to the *IPC* classification - since these three sectors tend to be more

---

<sup>8</sup> International Patent Classification

<sup>9</sup> The technological specialization index is designed to control for the existence of Marshall-Arrow-Romer (*MAR*) externalities derived from the regional specialization in certain technological sectors. To calculate this, we employ the following formula

$$SpIn_i = \frac{1}{2} \sum \left| \frac{PAT_{ij}}{PAT_i} - \frac{PAT_{Cj}}{PAT_C} \right|, \quad (f.1)$$

where *PAT* is the number of patents in each region  $i$  for each sector  $j$ , expressed as a difference for the whole sample ( $C$ ). Thus, a positive and significant coefficient for this variable would point to the existence of *MAR* externalities.

research intensive, we need to control for their presence in each region.<sup>10</sup> A summary of the variable descriptions and the data sources can be found in Appendix 2.<sup>11</sup>

The empirical analysis is conducted for the NUTS 2 regions of 17 Western European countries (EU-15 plus Norway and Switzerland) –see Appendix 1. However, NUTS 1 regions have also been considered for the UK, and NUTS 0 for Denmark and Switzerland, due to data constraints.<sup>12</sup>

## 5. Preliminary results

Table 1 presents the results of the OLS estimations with robust standard errors to control for heteroskedasticity of unknown form. In columns (i), (ii), and (iii) the estimations are presented without country dummies. It should be borne in mind that we have estimated three specifications using alternative proxies for the network scale. In principle, the coefficients can be interpreted as elasticities, since the variables in the regression are expressed in natural logarithmic form: the proportional increase in patenting activity in response to a 1% increase in a given explanatory variable. In this way, the interpretation of R&D stock and human capital coefficients is straightforward. In the case of the remaining coefficients, some doubts must remain regarding their direct interpretation. Since the functional relationship between these variables and their response is decided in an ad-hoc fashion, their sign and significance can be fairly informative, although any interpretation of their magnitude should be treated with caution (BODE, 2004).

Some specific results are worth highlighting. The R&D stock and human capital parameters are, in general, strongly significant, with the expected sign, and with similar values to those reported elsewhere in the literature. The control variables tend to be significant and with the expected sign, although there are a number of exceptions. The specialisation index is only significant (and positive) when country dummies are included – the same is true in the subsequent tables. The share of patents in organic chemistry has a positive impact on innovation, whilst, surprisingly, biotechnology and pharmaceutical shares correlate negatively – the latter, always significantly. We believe that manufacturing employment (strongly positive and significant) must capture part of the effect of these three sectors on innovation. Strikingly, population density is significant, but contrary to our

---

<sup>10</sup> Although overall employment in these sectors would be a better proxy, these data are not available.

<sup>11</sup> We added a small value, 0.0000001, to all the explanatory variables presenting zero values in at least one observation to allow for a logarithmic transformation. Three Greek regions filed no patent applications in our period of analysis and so were removed from the study.

<sup>12</sup> We also consider the German land of Sachsen-Anhalt as a single NUTS 1 region, and we have omitted the regions of Las Canarias, Ceuta, Melilla, Madeira, Açores, Guadeloupe, Martinique, Guyane and Reunion due to their distance from continental Europe. Moreover, the two NUTS2 regions of Trentino-Alto Adige/Südtirol (Italy) are considered as a single region.



predictions, it is negative. This can be interpreted in several ways: first, the positive effects of market potential might already be taken into account with the inclusion of GVA per capita, which is always significant and positive. Moreover, it might be the case that overly dense areas may suffer several costs related to congestion effects of agglomeration (negative externalities). Alternatively, the result may also reflect the fact that the regions in our sample are particularly heterogeneous. Finally, the significant yet negative effect of population density on innovation might point to the absence of urbanisation (inter-industry) economies, at the expense of specialisation economies.

The focus variables of this study are also significant. Mobility, for example, is always significant at 1%, presenting parameters around 0.26 and 0.29, whilst the relationship between the scale of the networks and innovation is always positive and strongly significant –except in the case of the size of the largest component variable. In addition, network density shows a negative impact on innovation intensity, which bestows credibility to GRANOVETTER's (1985) arguments about weak ties and innovation. Encouragingly, our results are broadly similar to the findings reported for the US, although the approach we adopt is slightly different.

[Insert Table 1 about here]

In short, the empirical analyses undertaken here support our hypotheses concerning the importance of labour mobility and networks in the local labour market for the diffusion and creation of regional innovations. However, several extensions to this initial approach can now be made.

## 6. Extensions

In this section, several extensions to the relationships across regions are included through the formulation of three additional hypotheses.

### **Spatial structure of innovation data**

*H3. The spatial structure of the data also matters.*

Although included as an additional hypothesis, it is already a stylised fact that innovation and knowledge variables are, by large, spatially correlated, irrespective of the level of regional disaggregation, the time span, or the sample of regions analysed (ACS et al., 2002, ANSELIN et al.,

1997; AUTANT-BERNARD and LESAGE, 2009; BODE, 2004; MORENO et al., 2005b). Indeed, knowledge tends to flow locally, but there is no reason to assume that knowledge stops flowing because of regional borders (AUDRETSCH and FELDMAN, 2004). Against this background, here we are interested in elucidating whether our preliminary findings are robust to the correct specification of these spatial effects. To step in that direction, we now assume that the R&D productivity index is also a function of the innovation intensity of the neighbouring regions of a given region,  $\prod_{\substack{j=1 \\ j \neq i}}^N y_j^{\rho_y w_{y,ij}}$ ,

where  $w_{y,ij}$  are bilateral spatial weights, which will depend on the distance between regions  $i$  and  $j$ , while  $\rho_y$  is the output elasticity of cross-regional effects of innovation intensity.

$$A_i = g(A_C, \prod_{\substack{j=1 \\ j \neq i}}^N y_j^{\rho_y w_{y,ij}}, MOB_i, NET_i, DENS_i, Z_i, X_i). \quad (10)$$

Expressing this in log form in the multiplicative version of 10 and 3, and making the necessary rearrangements, we obtain the following model for estimation, including once again a random error term,  $\varepsilon_i$ :

$$\begin{aligned} \ln PATpc_{it} = & \xi + \rho W \ln y_j + \beta \cdot \ln k_{it-1} + \alpha \cdot \ln l_{it-1} + \omega_1 \cdot \ln MOB_{it-1} + \\ & + \omega_2 \cdot \ln NET_{it-1} + \omega_3 \cdot \ln DENS_{it-1} + \omega_j \cdot \ln Z_{it-1} + \omega_n \cdot \ln X_{it-1} + \quad , \quad (11) \\ & + \omega_N \cdot \ln DUMMY\_COUNTRY\_EFFECT + \varepsilon_i \end{aligned}$$

where

$$W = \sum_{j=1}^N w_{ij} . \quad (12)$$

Moran's I tests and Lagrange Multiplier (LM) tests (robust and non robust) are also calculated at the end of Table 1. As can be seen, OLS residuals remain spatially correlated after the estimations – even when country dummies are included, although quantitatively their importance is relatively small. Moreover, LM tests seem to point to the need to estimate the spatial lag model, as was suggested in 10. Furthermore, and in close relation to issues affecting innovation, the spatial lag model suggests an economically meaningful interpretation, i.e., spillovers across regions captured in the spatial lag of the dependent variable; whilst the spatial error autocorrelation could be attributed to noise (ANSELIN, 1988) or even to a misspecification of the model (ANDERSSON and GRÅSJÖ,

2009). However, as this is somewhat unclear, we chose to check for any remaining spatial correlation in the residuals following the estimation of this model.

As far as the spatial econometric techniques are concerned, the inclusion of the spatial lag of the dependent variable as an additional regressor introduces endogeneity and, therefore, OLS estimations are no longer consistent. Thus, such models are usually estimated by maximum likelihood, though this method requires specific assumptions regarding the distribution and properties of the variables (KELEJIAN and PRUCHA, 1998). Moreover, the possible estimation of the spatial lag and error model simultaneously will make the use of the technique cumbersome, if not impossible (FINGLETON and LE GALLO, 2008). In such a setting, spatial two stage least squares (S2SLS) procedures are preferred, where the instruments are the spatial lags of the remaining explanatory variables.<sup>13</sup>

A crucial decision in spatial econometrics is the choice of the weighting matrix to characterize the spatial relationships between regions. Different matrices can lead to different conclusions (HARRIS and KRAVSTOVA, 2009), and it is therefore essential to ascertain which features illustrate the linkages across locations. In our paper we expressly specify a geographical weights matrix, although we are aware of the criticism such matrices have received from a number of scholars who claim they are oversimplified (BOSCHMA, 2005).

Nonetheless, as we see below, we are interested precisely in identifying what part is attributable to geographical space when other more meaningful variables describing interactions across regions are considered. However, instead of simply using contiguity matrices, we apply the concept of k-nearest neighbours. Thus, we measure the Euclidean distance to a fixed number of neighbours, k. In our case, we perform our analysis using a matrix where k=5, because both the mean and the median number of neighbours for the contiguity matrix is located around 5.

Table 2 shows the results of the spatial lag model. As can be seen, the parameter for the spatial lag is positive and strongly significant, pointing to the existence of relevant knowledge flows across regions affecting their respective innovation outcomes. Again, this result is in line with findings reported elsewhere. If we compare these results for the remaining explanatory variables with the previous OLS estimations, we see that both the value of the parameters and their significance remain practically unchanged.

[Insert Table 2 about here]

---

<sup>13</sup> We use the first and second order spatial lags, as suggested in KELEJIAN et al. (2004).

As can be seen at the bottom of Table 2, the modified Moran's I test for 2SLS suggested in ANSELIN and KELEJIAN (1997) points to the existence of a residual correlation –albeit not very strong - at least in three of the estimations (iv, v, and vi). Given these results, we believe that spatial correlation is no longer a problem, yet nevertheless we seek to confirm this below. We address any remaining concerns by implementing recent developments in spatial econometrics by KELEJIAN and PRUCHA (2007) which, additionally and contrary to earlier work, do not impose a specific functional form of the spatial correlation of the error term<sup>14</sup>, i.e. the spatial HAC estimator of the V-C matrix. The rationale behind this technique is drawn from the time-series results, and basically it is a non-parametric technique for estimating the V-C matrix using the averages of cross-products of residuals, the range of which is determined by a kernel function. This kernel function takes the form of  $K(d_{ij}/d)$ , with  $d_{ij}$  the distance between regions i and j, and  $d$  the bandwidth -  $K(d_{ij}/d)$  equals 0 when  $d_{ij} \geq d$ . Here, we use the triangular kernel, i.e.,  $K(d_{ij}/d) = 1 - (d_{ij}/d)$ . We perform a number of robustness checks with Epanechnikov and bisquare kernels - available upon request - which do not make any great changes to the results. We use a variable bandwidth with Euclidean distances to the ten nearest neighbours. Estimations using other distances or a different number of neighbours do not alter our results.

Table 3 summarises the results of the estimation of the spatial model with a spatial lag and the spatial heteroskedastic and autocorrelation consistent estimation of the V-C matrix, using triangular kernels and the distances to the 10<sup>th</sup> nearest neighbours. Encouragingly, the significance of the parameters in Table 2 remains virtually unchanged.

[Insert Table 3 about here]

### **Networks and mobility: Cross-regional linkages**

*H4. Cross-regional collaborations and inter-regional mobility enhance creativity. The more inventors collaborate with fellow inventors outside the region, the greater are the returns on innovation. Similarly, the greater the number of inventors moving into a given region, the greater is the patenting activity of that region.*

---

<sup>14</sup> Although the inclusion of a spatial lag of the dependent variable to summarise a broader set of externalities is theoretically straightforward, the a priori functional form of the spatial process in the disturbance term is less clear and that is why we find the validity of the approach proposed by K-P (2007) more convincing here.

Thus, in Table 4 we include the variable `OUT_COLL`, which compares the number of co-inventors outside the region with the total number of inventors within a region. We also include the variable `Inward Migration Rate (IMR)`, which is calculated counting the number of inventors moving into the region compared to the total number of inventors identified in that region. We expect, other things being equal, a positive and significant effect for both variables, for the same reason that a positive effect was expected for the scale of networks and labour mobility within the region. This cross-regional knowledge should be critical for innovation because of the fact that the imported knowledge is even more relevant than that already held locally. Here, we only show the estimations with the `CONNECTIVITY` variable, though the other models do not change our conclusions greatly.

The results (see Table 4 -columns i-vi) corroborate the importance of outside collaborators which, even though it presents smaller parameters than those presented by other networking variables, is significant in the specifications that include country dummies. The same results are not found in the case of the `IMR`, where positive parameters are only slightly significant in the OLS estimation without country dummies, and they are not significant for any of the other specifications. We interpret these results as follows: the geographical mobility of inventors is quite localised (BRESCHI and LISSONI, 2009), and the aggregate unit of analysis (NUTS2, and some NUTS1 and NUTS0) fails to identify a number of movements that occur within the region and between, for instance, NUTS3 regions. Networks, however, can connect points further apart in space (PONDS et al., 2009), and thus help to overcome geographic distances (SINGH, 2005). Therefore, networks may be useful for transferring knowledge over longer distances. In addition, it might also be argued that the actual origin of these inventors - whether or not they originate from a more innovative region - also matters.

*H5. The strength of the spatial autocorrelation of the data falls markedly when specific variables of inter-regional linkages are included in the specification.*

Finally, we examine a somewhat more challenging hypothesis. We strongly believe that most of the knowledge flows across regions in the knowledge production process are driven by networks of research collaborations and labour mobility. Therefore, we expect the values of the spatial autocorrelation tests to fall significantly when these two variables are included in the regressions. We think that this will be particularly true for the case of the residual spatial correlation following the estimation of the spatial lag model. If this is indeed the case, then we have found a way to model interdependences across regions in the production of knowledge that does not include the spatial error model - which is, in fact, a kind of ad-hoc solution for a misspecified model.

If we compare the spatial residual tests at the bottom of Table 4 with those at the bottom of Tables 1 and 2, we are unfortunately unable to assert categorically that the inclusion of economically meaningful linkages across regions has any significant effect on the residuals in our previous estimations. In short, we can observe a reduction in the OLS spatial residuals tests (especially those of the spatial LM and RLM) if we compare Table 1 (columns i and iv) and Table 4 (columns i and ii). However, if we compare Table 2 (columns i and iv) and Table 4 (columns iii and iv), we do not observe any reduction in the value of the Moran's I. Consequently, we should recognise that the inclusion of these variables does not eliminate completely the spatial correlation in the residuals, so the hypothesis cannot be confirmed (although neither can it be rejected). Further research on this question is, therefore, required.

[Insert Table 4 about here]

## 7. Conclusions, implications and limitations

The research conducted here sought to assess the importance of market-based knowledge flow mechanisms, namely networks of co-invention and labour mobility, on regional innovation, as opposed to the impact from R&D efforts or pure externalities. Within a KPF framework, several hypotheses have been suggested and, although we are unable to confirm them all, a number of interesting conclusions can be identified.

We find strong support for the positive relationship between regional labour market mobility and regional innovation intensity. The influence of networks is also fairly important, but the strength of these ties (measured as the network density) was found to have a negative influence on innovation. In line with studies elsewhere, we rely on the explanations proffered by GRANNOVETTER (1985) concerning the importance of weak ties for innovation. This idea is further strengthened with the inclusion in our regressions of a proxy for the number of collaborations with inventors outside the region (positive and strongly significant), which leads us to conclude that weak, distant ties are indeed more important for innovation.

Contrary to our predictions, however, the inflow of inventors from other regions does not have a markedly significant impact on innovation outcomes. In principle, we attribute these findings to the

level of regional disaggregation at which our analysis was conducted, although, in line with SHALEM and TRAJTENBERG (2008), the existence of certain short-term costs of a change in location should not be overlooked.

However, the results reported herein reveal certain tendencies and a number of policy implications can be directly derived from them. In the case of the network variable, it seems advisable, both from the viewpoint of the firm and from that of the region, to promote such collaborations so as to acquire as much external knowledge as possible (see similar claims in SINGH, 2005). Additionally, given the results for such variables as network density and outside collaborations, it would seem that the promotion of inter-firm collaborations is preferable to that of intra-firm collaborations, whilst inter-regional collaborations should receive greater attention than intra-regional linkages. Policy recommendations regarding mobility within the local labour market are not so straightforward. Although mobility seems to be desirable at an aggregate level, and also at the inventor level, it could be understood as a zero-sum game for firms. A policy option could, therefore, be to promote the competition for talent at the inter-regional and, in particular, the international levels.

It should be borne in mind, however, that our approach presents several limitations. Below we highlight some of the drawbacks to our analysis, and should serve as the basis for improved approaches.

Patenting activity does not take into account all the research collaboration relationships that exist between individuals or all the mobility patterns of those individuals. In so far as other variables capturing similar concepts could be assembled, the analysis conducted should be repeated. Related to the first of these shortcomings, extending our sample to include all EPO patents and not just those that had been filed to the PCT ought to shed more light on the relationships analysed herein.

As regards the raw data used here, improvements to the suggested algorithms (see MIGUÉLEZ and MIGUÉLEZ, 2010) should certainly result in variables with fewer measurement errors and, ultimately, to more consistent estimations. Moreover, even though outside collaborators are included here as a regressor, our method of measuring regional networks imposes implicitly artificial boundaries on the administrative limits of the region, thereby failing to identify possible links between inventors when they are linked via a third inventor located outside the region. Suitable variables able to capture this phenomenon might alter some of the results.

Next, as regards the KPF framework: a theoretical model of spatial interactions occurring across regions in the production of knowledge, and which the estimated model is based on, is lacking in this



paper and, arguably, throughout the literature of spatial econometrics and knowledge spillovers. Analogously, a specific, more meaningful modelling of these interactions in the spatial weights matrix is needed. Fortunately, co-patenting and mobility data are opening up a promising line of research in this direction, to which we will no doubt focus our attention sooner rather than later. Moreover, lagging variables of the right hand side of the models seeks to reduce the problems of endogeneity and reverse causality, thus implicitly assuming that weak exogeneity is sufficient to obtain consistent results. However, we ought to acknowledge the possibility that patenting activity in earlier years might well have influenced the levels of certain independent variables, and therefore consistency will be affected. To the best of our knowledge, suitable instruments have still to be found for the explanatory variables, and so further research along these lines must first be undertaken.<sup>15</sup>

---

<sup>15</sup> An earlier version of this article sought to deal with endogeneity problems by instrumenting the focus variables using the three-group procedure suggested by KENNEDY (1992). The method is thought to cope with measurement errors, and the results of the procedure have to be treated with caution. The three-group method involves sorting the variables into three groups of equal size, taking the value 1 if the observation is in the highest third of the variable, 0 if it is in the middle, and -1 if the value is in the lowest third of the regressor. The results – available upon request – make few changes to our conclusions, with the exception of “outside collaborators” variable, which was no longer significant at 10%.

## References

- ACS Z., ANSELIN L. and VARGA A. (2002) Patents and innovation counts as measures of regional production of new knowledge, *Research Policy* 31, 1069-85;
- AGRAWAL A., COCKBURN I. and MCHALE J. (2006) Gone but not forgotten: labour flows, knowledge spillovers, and enduring social capital, *Journal of Economic Geography* 6, 571-91;
- AKÇOMAK I. S. and TER WEEL B. (2009) Social capital, innovation and growth: Evidence from Europe, *European Economic Review* 53, 544-67;
- ALMEIDA P. and KOGUT B. (1999) Localisation of knowledge and the mobility of engineers in regional networks, *Management Science* 45, 905-17;
- ANDERSSON M. and GRÅSJÖ U. (2009) Spatial dependence and the representation of space in empirical models, *The Annals of Regional Science* 43, 159-80;
- ANSELIN, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, The Netherlands.
- ANSELIN L., VARGA A. and ACS Z. (1997) Local Geographic Spillovers between University Research and High Technology Innovations, *Journal of Urban Economics* 42, 422-48;
- ANSELIN L. and KELEJIAN H. H. (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors, *International Regional Science Review* 20, 153-82;
- AUDRETSCH D. B. and FELDMAN, M. (2004) Knowledge Spillovers and the Geography of Innovation, in V. HENDERSON and J. THISSE (Eds) *Handbook of Urban and regional Economics, Volume 4*;
- AUTANT-BERNARD C., MAIRESSE J. and MASSARD N. (2009) Spatial knowledge diffusion through collaborative networks. Introduction, *Papers in Regional Science* 86, 3, 341-50;
- AUTANT-BERNARD C. and LESAGE J. (2009) Quantifying Knowledge Spillovers using Spatial Econometric Models, working paper;
- BALCONI M., BRESCHI S. and LISSONI F. (2004) Networks of inventors and the role of academia: an exploration of Italian patent data, *Research Policy* 33, 127-45;
- BERGMAN E. M. (2009) Embedding network analysis in spatial studies of innovation, *The Annals of Regional Science* 43, 3, 559-65;
- BETTENCOURT L., LOBO J., HELBING D., KÜHNERT C., and WEST G. (2007a) Growth, innovation, scaling, and the pace of life in cities, *Proceeding of the National Academy of Sciences* 104, 7301-06;
- BETTENCOURT L., LOBO J., and STRUMSKY D. (2007b) Invention in the city: Increasing returns topatenting as a scaling function of metropolitan, *Research Policy* 36, 107-20;
- BODE E. (2004) The spatial pattern of localized R&D spillovers: an empirical investigation for Germany, *Journal of Economic Geography* 4, 43-64;
- BOSCHMA R. (2005) Proximity and Innovation: A Critical Assessment, *Regional Studies*, 39.1, 61-74;

- BOSCHMA R., ERIKSSON R. and LINDGREN U. (2009) How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity, *Journal of Economic Geography* 9, 169-90;
- BOTTAZZI L. and PERI G. (2003) Innovation and spillovers in regions: Evidence from European patent data, *European Economic Review* 47, 687 - 710;
- BRESCHI S. and CATALINI C. (2009) Tracing the linkages between science and technology: An exploratory analysis of the research networks among scientists and inventors, submitted to *Research Policy*;
- BRESCHI S. and LISSONI F. (2004) Knowledge networks from patent data: methodological issues and research targets, in MOED H., GLÄNZEL W. and SCHMOCH U. (Eds) *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*, pp. 613-643. Springer Verlag, Berlin.
- BRESCHI S. and LISSONI F. (2006) Cross-firm inventors and social networks: localised knowledge spillovers revisited, *Annales d'Economie et de Statistique*, 79-80;
- 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, 4, 439-68;
- CORREDOIRA R. A. and ROSENKOPF L. (2006) Learning from those who left: the reverse transfer of knowledge through mobility ties, *Management Department Working Paper*
- COWAN R. and JONARD N. (2004) Network structure and the diffusion of knowledge, *Journal of Economic Dynamics and Control* 28, 1557-75;
- CRESPI G., GEUNA A. and NESTA L. (2007) The mobility of university inventors in Europe, *Journal of Technology Transfer* 32, 3, 195-215;
- DE CLERCQ D. and DAKHLI M. (2004) Human capital, social capital, and innovation: A multi-country study, *Entrepreneurship and Regional Development* 16, 2, 107-28;
- DE DOMINICI L., FLORAX R. J. G. M. and DE GROOT H. L. F. (2007) Regional clusters of economic activity in Europe: Are social and geographical proximity the key determinants?, *Knowledge for Growth - Industrial Research & Innovation (IRI)*
- DÖRING, T. and SCHNELLENBACH, J. (2006) What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature, *Regional Studies* 40.3, 375-395;
- FINGLETON B. and LE GALLO J. (2008) Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties, *Papers in Regional Science* 87, 319-39;
- FLEMING L., KING C., and JUDA A. (2007) Small worlds and innovation, *Organization Science* 14, 5, 375-93;
- GOMES-CASSERES B., HAGEDOORN J., and JAFFE A. B. (2006) Do alliances promote knowledge flows?, *Journal of Financial Economics* 80, 5-33;
- GRANOVETTER M. S. (1985) The strength of weak ties, *American Journal of Sociology* 78, 1360-80;

- GRILICHES Z. (1979) Issues in assessing the contribution of research and development to productivity growth, *The Bell Journal of Economics* 10, 1, 92-116;
- GRILICHES Z. (1991) Patent Statistics as Economic Indicators: A Survey, NBER Working Papers No. 3301, National Bureau of Economic Research;
- GROSSMAN. G. M. and E. HELPMAN (1991) *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, MA.
- HARRIS R. and KRAVTSOVA V. (2009) In Search of 'W', SERC Discussion Paper 17, London;
- HAUSMAN J.A., HALL B.H. and GRILICHES Z. (1984) Econometric models for count data with an application to the patent-R&D relationship, *Econometrica* 51, 4, 909-38;
- HOEKMAN J., FRENKEN K. and VAN OORT F. (2009) The geography of collaborative knowledge production in Europe, *The Annals of Regional Science* 43, 3, 721-38;
- HOISL K. (2007) Does mobility increase the productivity of inventors?, *Journal of Technology Transfer* 34, 212-25;
- HOISL K. (2009) Tracing mobile inventors: The causality between inventor mobility and inventor productivity, *Research Policy* 36, 5, 615-36;
- JAFFE A. B. (1986) Technological opportunity and spillovers of R&D: evidence from firms patents, profits and market value, *American Economic Review* 76, 5, 984-1001;
- JAFFE A. B. (1989) Real effects of academic research, *American Economic Review* 79, 5, 957-70;
- KELEJIAN, H. H. and PRUCHA, I. R. (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics* 17, 99-21;
- KELEJIAN H. H. and PRUCHA I. R. (2007) HAC estimation in a spatial framework, *Journal of Econometrics* 140, 131-54;
- KELEJIAN H. H., PRUCHA I. and YUZEFOVICH Y. (2004) Instrumental Variable Estimation of a Spatial Autorregressive Model with Autorregressive Disturbances: Large and Small Sample Results, in J. LESAGE and R. K. PACE (Eds) *Spatial and Spatiotemporal Econometrics*, *Advances in Econometrics*, Vol. 18, pp. 163-198. Elsevier, New York.
- KENNEDY, P. (1992) *A guide to econometrics*, 3<sup>rd</sup> Edition. Blackwell, Oxford.
- KIM J., LEE S. J. and MARSCHKE G. (2006) International Knowledge Flows: Evidence from an Inventor-Firm Matched Data Set, NBER Working Paper No. 12692
- KROLL H. (2009) Spillovers and proximity in perspective: A network approach to improving the operationalisation of proximity, *Working Papers Firms and Regions* No. R2/2009, Fraunhofer ISI;
- LENZI C. (2009) Patterns and determinants of skilled workers' mobility: evidence from a survey of Italian inventors, *Economics of Innovation and New Technology* 18, 2, 161-79;
- LOBO J. and STRUMSKY D. (2008) Metropolitan patenting, inventor agglomeration and social networks: A tale of two effects, *Journal of Urban Economics* 63, 871-84;

- LUCAS R. E. (1988) On the mechanics of economic development, *Journal of Monetary Economics* 22, 3-42;
- MAGGIONI M. A. and UBERTI T. E. (2008) Knowledge networks across Europe: which distance matters?, *The Annals of Regional Science* 43, 3, 691-720;
- MIGUÉLEZ E. and MIGUÉLEZ I. G. (2010) Singling out individual inventors from patent data, unpublished manuscript,  
[http://ub.academia.edu/documents/0058/3908/Miguelez\\_Miguelez\\_2010\\_-\\_Singling\\_out\\_individual\\_inventors\\_from\\_patent\\_data.pdf](http://ub.academia.edu/documents/0058/3908/Miguelez_Miguelez_2010_-_Singling_out_individual_inventors_from_patent_data.pdf);
- MIGUÉLEZ E., MORENO R. and ARTÍS M. (2009) Does social capital reinforce technological inputs in the creation of knowledge? Evidence from the Spanish regions, forthcoming in *Regional Studies*
- MORENO R., PACI R. and USAI S. (2005a) Spatial spillovers and innovation activity in European regions, *Environment and Planning A* 37, 10, 1793-812;
- MORENO R., PACI R. and USAI S. (2005b) Geographical and sectoral clusters of innovation Europe, *Annals of Regional Science* 39, 715-39;
- PERI G. (2005) Determinants of Knowledge Flows and Their Effect on Innovation, *The Review of Economics and Statistics* 8, 2, 308-22;
- PONDS R., VAN OORT F. and FRENKEN K. (2007) The geographical and institutional proximity of research collaboration, *Papers in Regional Science* 86, 3, 423-43;
- PONDS R., VAN OORT F. and FRENKEN K. (2009) Innovation, spillovers and university-industry collaboration: an extended knowledge production function approach, *Journal of Economic Geography*, doi:10.1093/jeg/lbp036;
- RODRÍGUEZ-POSE A. and CRESCENZI R. (2008) Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe, *Regional Studies* 42, 1, 51-67;
- ROMER P. M. (1986) Increasing Returns and Long-run Growth, *Journal of Political Economy* 94, 5, 1002-37;
- ROMER P. M. (1990) Endogenous Technological Change; *Journal of Political Economy* 98, 5, 71-102;
- ROSENKOPF L. and ALMEIDA P. (2003) Overcoming Local Search Through Alliances and Mobility, *Management Science* 49, 6, 751-66;
- SAXENIAN A. (1994) *Regional Advantage. Culture and Competition in Silicon Valley and Route 128.* Harvard University Press, Cambridge, MA.
- SINGH J. (2005) Collaborative Networks as Determinants of Knowledge Diffusion Patterns, *Management Science* 51, 5, 756-70;
- SINGH J. AND AGRAWAL A. K. (2009) Recruiting for Ideas: A Difference-in-Differences Approach for Estimating the Effect of Mobility on Access to an Inventor's Prior Knowledge, INSEAD Working Paper No. 2009/46/ST;
- SHALEM R. and TRAJTENBERG M. (2008) Software patents, inventors and mobility, Working Paper

SONG J., ALMEIDA P. and WU G. (2003) Learning-by-hiring: When is mobility more to facilitate interfirm knowledge transfer?, *Management Science* 49, 4, 351-65;

TER WAL A. L. J. and BOSCHMA R. (2009) Applying social network analysis in economic geography: framing some key analytic issues, *The Annals of Regional Science* 43, 739-56;

TRIPPL M. and MAIER G. (2007) Knowledge spillover agents and regional development, *SRE-Discussion* 2007/01

THOMA G. and TORRISI S. (2007) Creating Powerful Indicators for Innovation Studies with Approximate Matching Algorithms. A test based on PATSTAT and Amadeus databases, *CESPRI Working Papers*, No. 211;

THOMA G., TORRISI S., GAMBARDELLA A., GUELLEC D., HALL B.H., and HARHOFF D. (2009a) Methods and software for the harmonization and combination of datasets: A test based on IP-related data and accounting databases with a large panel of companies at the worldwide level, mimeo

THOMA G., TORRISI S., GAMBARDELLA A., GUELLEC D., HALL B.H., and HARHOFF D. (2009b) Harmonizing and Combining Large Datasets: An Application to Firm-level Patent and Accounting Data, *STI Working Paper*, OECD, Paris.

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

## Appendices

### Appendix 1: List of countries (and number of regions in each one)

Austria -AT- (9), Belgium -BE- (3), Switzerland -CH- (1), Germany -DE- (39), Denmark -DK- (1), Spain -ES- (16), Finland -FI- (5), France -FR- (22), Greece -GR- (10 (3 regions were dropped)), Ireland -IE- (2), Italy -IT- (20), Luxemburg -LU- (1), the Netherlands -NL- (12), Norway -NO- (7), Portugal -PT- (5), Sweden -SE- (8), United Kingdom -UK- (12).

### Appendix 2: Variables, data construction, and data source

Variable	Proxy	Dates	Source
Patents per capita	Patents per million inhabitants	2001-2003	REGPAT and Eurostat
R&D stock p.c.	Stock of R&D p.c. (euros per 1000 inhabitants). Perpetual inventory method using R&D expenditure	1996-2000	Eurostat
Human capital	HRST (Occupations) to total population	1993-2002	Eurostat
Mobility	Average no. assignees per inventor	1993-2002	REGPAT and OECD HAN
Connectivity	Absolute number of edges	1993-2002	REGPAT
Size of largest component	% inventors in the largest component	1993-2002	REGPAT
Edges per inventor	Number of edges to no. of inventors	1993-2002	REGPAT
Network density	$DENS_i = \frac{T}{Q(Q-1)/2}$	1993-2002	REGPAT
GVA per capita	Gross value added per capita	1996-2000	Eurostat
Population density	Population over total area (km <sup>2</sup> )	1996-2000	Eurostat
Specialisation Index	$SpIn_{it-1} = \frac{1}{2} \sum \left  \frac{PAT_{ij}}{PAT_i} - \frac{PAT_{Nj}}{PAT_N} \right $	1996-2000	REGPAT
Technological firms	No. of assignees to manufacturing establishments	1996-2000	REGPAT and Eurostat
Manufacturing	Share of manufacturing employment	1996-2000	Eurostat
% Organic chemistry	Share of patents in IPC chemistry	1996-2000	REGPAT
% Pharmaceuticals	Share of patents in IPC pharmaceuticals	1996-2000	REGPAT
% Biotechnology	Share of patents in IPC biotechnology	1996-2000	REGPAT
Outside collaborators	No. of inventors outside a region co-authoring patents with inventors inside the region	1993-2002	REGPAT
Inward Migration Rate	No. of inflows to no. of inventors	1993-2002	REGPAT



**Table 1. OLS estimations with White robust standard errors. Dep. Var.: Patents pc.**

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	OLS	OLS	OLS	OLS	OLS	OLS
Intercept	-5.25*** (0.63)	-5.44*** (0.73)	-5.20*** (0.63)	-3.43*** (0.99)	-3.13*** (1.10)	-3.37*** (0.99)
ln(STOCK_R&Dpc)	0.26*** (0.06)	0.39*** (0.07)	0.26*** (0.06)	0.16*** (0.05)	0.29*** (0.08)	0.17*** (0.05)
ln(HRST)	0.53** (0.23)	0.58** (0.25)	0.52** (0.23)	0.75** (0.32)	0.92** (0.39)	0.74** (0.32)
ln(MOB)	0.26*** (0.03)	0.29*** (0.03)	0.26*** (0.03)	0.27*** (0.03)	0.29*** (0.04)	0.27*** (0.03)
ln(DENS)	-0.16*** (0.04)	-0.05* (0.03)	-0.24*** (0.06)	-0.21*** (0.03)	-0.06 (0.04)	-0.33*** (0.05)
ln(CONECTIVITY)	0.08*** (0.02)			0.12*** (0.02)		
ln(SIZE_LC)		0.03 (0.04)			0.08* (0.04)	
ln(TIES_INV)			0.16*** (0.04)			0.22*** (0.05)
ln(POP_DENS)	-0.09*** (0.03)	-0.06* (0.03)	-0.09*** (0.03)	-0.15*** (0.04)	-0.08** (0.04)	-0.15*** (0.04)
ln(GVApc)	1.08*** (0.18)	1.06*** (0.19)	1.08*** (0.18)	0.65*** (0.18)	0.49** (0.23)	0.66*** (0.18)
ln(TECH_FIRMS)	0.14*** (0.04)	0.15*** (0.04)	0.14*** (0.04)	0.39*** (0.11)	0.42*** (0.12)	0.39*** (0.11)
ln(SPIN)	0.05 (0.13)	-0.16 (0.14)	0.05 (0.13)	0.26** (0.11)	0.02 (0.12)	0.26** (0.11)
ln(BIOTECH)	-0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)	0.01 (0.01)	0.01 (0.02)	0.01 (0.01)
ln(PHARMA)	-0.06*** (0.02)	-0.05** (0.02)	-0.06*** (0.02)	-0.05** (0.02)	-0.04 (0.03)	-0.04** (0.02)
ln(CHEM)	0.04* (0.02)	0.05*** (0.02)	0.04* (0.02)	0.01 (0.02)	0.03* (0.02)	0.01 (0.00)
ln(MAN)	0.92*** (0.09)	1.04*** (0.09)	0.92*** (0.09)	0.85*** (0.08)	1.04*** (0.08)	0.85*** (0.08)
Country Dummies	no	no	no	yes	yes	yes
Moran's I	5.779	4.450	5.764	0.276	1.546	0.249
p-value	0.000	0.000	0.000	0.782	0.122	0.803
Spatial error LM	24.063	13.552	23.925	1.572	0.038	1.629
p-value	0.000	0.000	0.000	0.210	0.846	0.202
Spatial error RLM	4.563	0.250	4.535	9.198	4.719	9.261
p-value	0.033	0.617	0.033	0.002	0.030	0.002
Spatial lag LM	47.632	50.693	47.346	10.831	10.701	10.613
p-value	0.000	0.000	0.000	0.001	0.001	0.001
Spatial lag RLM	28.132	37.391	27.956	18.456	15.382	18.244
p-value	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted-R2	0.929	0.920	0.929	0.959	0.946	0.959
Sample size	173	173	173	173	173	173
Log-likelihood	-76.03	-86.26	-75.97	-19.06	-43.40	-19.18
AIC	180.07	200.59	179.94	92.11	140.79	92.37
Schwarz	224.21	244.66	224.08	177.25	225.93	177.51
Mean V.I.F.	3.44	2.74	4.10	5.87	5.39	6.29

Notes: Robust standard errors are presented in parentheses. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Moran's I tests of the residuals and spatial Lagrange Multiplier tests point to the existence of residual spatial autocorrelation. The mean of the Variance Inflation Factor is also presented, and collinearity does not pose a serious concern.

**Table 2. S2SLS Dep. Var.: Patents pc. Spatial lag model**

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	S2SLS	S2SLS	S2SLS	S2SLS	S2SLS	S2SLS
Intercept	-4.45*** (0.55)	-4.33*** (0.64)	-4.41*** (0.54)	-3.81*** (0.82)	-3.55*** (0.88)	-3.75*** (0.82)
W·lnPATpc	0.33*** (0.05)	0.38*** (0.05)	0.33*** (0.05)	0.26*** (0.07)	0.29*** (0.08)	0.26*** (0.07)
ln(STOCK_R&Dpc)	0.33*** (0.05)	0.42*** (0.06)	0.33*** (0.06)	0.20*** (0.05)	0.33*** (0.07)	0.21*** (0.05)
ln(HRST)	0.45*** (0.17)	0.47** (0.19)	0.44** (0.18)	0.81*** (0.27)	0.98*** (0.31)	0.81*** (0.27)
ln(MOB)	0.25*** (0.02)	0.27*** (0.02)	0.25*** (0.03)	0.27*** (0.02)	0.28*** (0.03)	0.27*** (0.02)
ln(DENS)	-0.14*** (0.03)	-0.05** (0.02)	-0.21*** (0.05)	-0.20*** (0.03)	-0.05* (0.03)	-0.31*** (0.05)
ln(CONECTIVITY)	0.07*** (0.02)			0.11*** (0.02)		
ln(SIZE_LC)		0.06* (0.04)			0.08** (0.04)	
ln(TIES_INV)			0.13*** (0.04)			0.21*** (0.04)
ln(POP_DENS)	-0.02 (0.03)	0.01 (0.03)	-0.02 (0.03)	-0.10*** (0.03)	-0.03 (0.03)	-0.10*** (0.03)
ln(GVApc)	0.40** (0.16)	0.29* (0.18)	0.40*** (0.16)	0.28* (0.17)	0.10 (0.19)	0.29* (0.17)
ln(TECH_FIRMS)	0.05 (0.03)	0.05 (0.04)	0.05 (0.03)	0.28*** (0.10)	0.29*** (0.10)	0.28*** (0.10)
ln(SPIN)	0.11 (0.11)	-0.08 (0.11)	0.11 (0.11)	0.26** (0.10)	0.03 (0.11)	0.26** (0.10)
ln(BIOTECH)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.02)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
ln(PHARMA)	-0.05*** (0.02)	-0.05** (0.02)	-0.05*** (0.02)	-0.05*** (0.02)	-0.04** (0.02)	-0.05*** (0.02)
ln(CHEM)	0.04** (0.02)	0.05*** (0.02)	0.04** (0.02)	0.01 (0.01)	0.04*** (0.01)	0.01 (0.01)
ln(MAN)	0.70*** (0.09)	0.77*** (0.09)	0.70*** (0.09)	0.68*** (0.09)	0.85*** (0.10)	0.68*** (0.09)
Country Dummies	no	no	no	yes	yes	yes
Pseudo-R2	0.939	0.935	0.939	0.961	0.953	0.961
Sample size	173	173	173	173	173	173
Moran's I	0.987	0.340	1.011	4.245	3.657	4.251
p-value	0.340	0.550	0.300	0.030	0.051	0.030
Log-likelihood	-51.01	-58.91	-51.12	-13.65	-40.52	-14.01
AIC	132.01	147.82	132.23	83.29	137.04	84.01
Schwarz	179.31	195.12	179.53	171.58	225.33	172.31

Notes: Robust standard errors are presented in parentheses. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Pseudo R2 is calculated as the ratio between the variance of the fitted value and the variance of the dependent variable. The Moran's I is the version suggested by ANSELIN and KELEJIAN (1997) which takes into account the existence of an endogenous regressor in the r.h.s. of the equation in the form of a spatial lag of the dependent variable. It is distributed as a Chi-square with 1 degree of freedom and presents significant spatial autocorrelation in models iv, v, and vi.

**Table 3. S2SLS Dep. Var.: Patents pc. Spatial HAC estimation V-C matrix**

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	SHAC	SHAC	SHAC	SHAC	SHAC	SHAC
Intercept	-4.45*** (0.54)	-4.33*** (0.61)	-4.41*** (0.55)	-3.81*** (0.68)	-3.55*** (0.74)	-3.75*** (0.68)
W·lnPATpc	0.33*** (0.05)	0.38*** (0.05)	0.33*** (0.05)	0.26*** (0.07)	0.29*** (0.09)	0.26*** (0.07)
ln(STOCK_R&Dpc)	0.33*** (0.06)	0.42*** (0.07)	0.33*** (0.05)	0.20*** (0.06)	0.33*** (0.08)	0.21*** (0.06)
ln(HRST)	0.45*** (0.17)	0.47** (0.18)	0.44** (0.17)	0.81*** (0.25)	0.98*** (0.32)	0.81*** (0.25)
ln(MOB)	0.25*** (0.03)	0.27*** (0.02)	0.25*** (0.02)	0.27*** (0.03)	0.28*** (0.03)	0.27*** (0.03)
ln(DENS)	-0.14*** (0.03)	-0.05** (0.02)	-0.21*** (0.05)	-0.20*** (0.03)	-0.05* (0.03)	-0.31*** (0.05)
ln(CONECTIVITY)	0.07*** (0.02)			0.11*** (0.02)		
ln(SIZE_LC)		0.06* (0.04)			0.08** (0.04)	
ln(TIES_INV)			0.13*** (0.04)			0.21*** (0.04)
ln(POP_DENS)	-0.02 (0.03)	0.01 (0.03)	-0.02 (0.03)	-0.10*** (0.03)	-0.03 (0.04)	-0.10*** (0.03)
ln(GVApc)	0.40** (0.17)	0.29 (0.18)	0.40*** (0.16)	0.28* (0.17)	0.10 (0.19)	0.29* (0.17)
ln(TECH_FIRMS)	0.05 (0.03)	0.05 (0.03)	0.05 (0.03)	0.28*** (0.09)	0.29*** (0.10)	0.28*** (0.09)
ln(SPIN)	0.11 (0.11)	-0.08 (0.11)	0.11 (0.11)	0.26** (0.10)	0.03 (0.10)	0.26** (0.10)
ln(BIOTECH)	-0.01 (0.02)	-0.01 (0.01)	-0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
ln(PHARMA)	-0.05*** (0.02)	-0.05** (0.02)	-0.05*** (0.02)	-0.05** (0.02)	-0.04** (0.02)	-0.05** (0.02)
ln(CHEM)	0.04** (0.02)	0.05*** (0.01)	0.04** (0.02)	0.01 (0.01)	0.04*** (0.01)	0.01 (0.01)
ln(MAN)	0.70*** (0.09)	0.77*** (0.09)	0.70*** (0.09)	0.68*** (0.09)	0.85*** (0.11)	0.68*** (0.09)
Country Dummies	no	no	no	yes	yes	yes
Pseudo-R2	0.939	0.935	0.939	0.961	0.953	0.961
Sample size	173	173	173	173	173	173
Log-likelihood	-51.01	-58.91	-51.12	-13.65	-40.52	-14.01
AIC	132.01	147.82	132.23	83.29	137.04	84.01
Schwarz	179.31	195.12	179.53	171.58	225.33	172.31

Notes: Robust standard errors are presented in parentheses. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Pseudo R2 is calculated as the ratio between the variance of the fitted value and the variance of the dependent variable.

**Table 4. OLS and S2SLS estimations. Includes outside collaborations and I.M.R.**

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	OLS	OLS	S2SLS	S2SLS	SHAC	SHAC
Intercept	-4.84*** (0.62)	-3.36*** (0.93)	-4.24*** (0.56)	-3.76*** (0.75)	-4.24*** (0.56)	-3.76*** (0.63)
W·lnPATpc			0.32*** (0.05)	0.26*** (0.07)	0.32*** (0.05)	0.26*** (0.07)
ln(STOCK_R&Dpc)	0.25*** (0.06)	0.15*** (0.05)	0.32*** (0.05)	0.19*** (0.05)	0.32*** (0.06)	0.19*** (0.05)
ln(HRST)	0.45* (0.23)	0.65** (0.30)	0.40** (0.17)	0.72*** (0.24)	0.40** (0.17)	0.72*** (0.24)
ln(MOB)	0.25*** (0.03)	0.27*** (0.03)	0.25*** (0.03)	0.27*** (0.02)	0.25*** (0.03)	0.27*** (0.03)
ln(IMR)	0.03** (0.01)	0.00 (0.01)	0.02 (0.01)	0.00 (0.01)	0.02 (0.01)	0.00 (0.01)
ln(DENS)	-0.16*** (0.04)	-0.22*** (0.04)	-0.14*** (0.03)	-0.21*** (0.03)	-0.14*** (0.03)	-0.21*** (0.03)
ln(CONECTIVITY)	0.08*** (0.02)	0.12*** (0.02)	0.07*** (0.02)	0.11*** (0.02)	0.07*** (0.02)	0.11*** (0.02)
ln(OUT_COLL)	0.02 (0.02)	0.03** (0.01)	0.02 (0.01)	0.03** (0.01)	0.02 (0.01)	0.03** (0.01)
ln(POP_DENS)	-0.08** (0.03)	-0.15*** (0.04)	-0.03 (0.03)	-0.11*** (0.03)	-0.03 (0.03)	-0.11*** (0.03)
ln(GVApc)	1.09*** (0.18)	0.75*** (0.16)	0.44*** (0.16)	0.38** (0.15)	0.44*** (0.16)	0.38** (0.15)
ln(TECH_FIRMS)	0.14*** (0.04)	0.39*** (0.10)	0.05 (0.03)	0.27*** (0.09)	0.05* (0.03)	0.27*** (0.08)
ln(SPIN)	0.10 (0.13)	0.28** (0.11)	0.14 (0.11)	0.28*** (0.10)	0.14 (0.11)	0.28*** (0.10)
ln(BIOTECH)	-0.01 (0.02)	0.01 (0.01)	-0.00 (0.01)	0.01 (0.01)	-0.00 (0.01)	0.01 (0.01)
ln(PHARMA)	-0.05*** (0.02)	-0.04* (0.02)	-0.05*** (0.02)	-0.04*** (0.02)	-0.05*** (0.02)	-0.04*** (0.02)
ln(CHEM)	0.03 (0.02)	0.01 (0.02)	0.03** (0.02)	0.01 (0.01)	0.03** (0.02)	0.01 (0.01)
ln(MAN)	0.91*** (0.09)	0.84*** (0.08)	0.70*** (0.09)	0.67*** (0.09)	0.70*** (0.09)	0.67*** (0.09)
Country Dummies	no	yes	no	yes	no	yes
Moran's I	5.416	0.912	1.625	4.303		
p-value	0.000	0.362	0.200	0.040		
Spatial error LM	21.053	0.504				
p-value	0.000	0.478				
Spatial error RLM	4.290	5.898				
p-value	0.038	0.015				
Spatial lag LM	41.062	10.935				
p-value	0.000	0.001				
Spatial lag RLM	24.299	16.329				
p-value	0.000	0.000				
Adjusted-R2	0.933	0.920	0.942	0.962	0.942	0.962
Sample size	173	173	173	173	173	173
Log-likelihood	-70.04	-86.26	-48.32	-11.27	-48.32	-11.27
AIC	172.09	200.59	130.64	82.53	130.64	82.53
Schwarz	222.54	244.66	184.24	177.13	184.24	177.13

Notes: Robust standard errors are presented in parentheses. Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Pseudo R2 is calculated as the ratio between the variance of the fitted value and the variance of the dependent variable. The Moran's I in iii and iv is the version suggested by ANSELIN and KELEJIAN (1997).

---

## ***Llista Document de Treball***

### *List Working Paper*

---

- WP 2010/01 “Research Networks and Inventors’ Mobility as Drivers of Innovation: Evidence from Europe” Miguélez, E. ; Moreno, R. ”
- WP 2009/26 “Social Preferences and Transport Policy: The case of US speed limits” Albalate, D.
- WP 2009/25 “Human Capital Spillovers Productivity and Regional Convergence in Spain” , Ramos, R ; Artis, M. ; Suriñach, J.
- WP 2009/24 “Human Capital and Regional Wage Gaps” ,López-Bazo,E. Motellón E.
- WP 2009/23 “Is Private Production of Public Services Cheaper than Public Production? A meta-regression analysis of solid waste and water services” Bel, G.; Fageda, X.; Warner. M.E.
- WP 2009/22 “Institutional Determinants of Military Spending” Bel, G., Elias-Moreno, F.
- WP 2009/21 “Fiscal Regime Shifts in Portugal” Afonso, A., Claeys, P., Sousa, R.M.
- WP 2009/20 “Health care utilization among immigrants and native-born populations in 11 European countries. Results from the Survey of Health, Ageing and Retirement in Europe” Solé-Auró, A., Guillén, M., Crimmins, E.M.
- WP 2009/19 “La efectividad de las políticas activas de mercado de trabajo para luchar contra el paro. La experiencia de Cataluña” Ramos, R., Suriñach, J., Artís, M.
- WP 2009/18 “Is the Wage Curve Formal or Informal? Evidence for Colombia” Ramos, R., Duque, J.C., Suriñach, J.
- WP 2009/17 “General Equilibrium Long-Run Determinants for Spanish FDI: A Spatial Panel Data Approach” Martínez-Martín, J.
- WP 2009/16 “Scientists on the move: tracing scientists’ mobility and its spatial distribution” Miguélez, E.; Moreno, R.; Suriñach, J.
- WP 2009/15 “The First Privatization Policy in a Democracy: Selling State-Owned Enterprises in 1948-1950 Puerto Rico” Bel, G.
- WP 2009/14 “Appropriate IPRs, Human Capital Composition and Economic Growth” Manca, F.
- WP 2009/13 “Human Capital Composition and Economic Growth at a Regional Level” Manca, F.
- WP 2009/12 “Technology Catching-up and the Role of Institutions” Manca, F.
- WP 2009/11 “A missing spatial link in institutional quality” Claeys, P.; Manca, F.
- WP 2009/10 “Tourism and Exports as a means of Growth” Cortés-Jiménez, I.; Pulina, M.; Riera i Prunera, C.; Artís, M.
- WP 2009/09 “Evidence on the role of ownership structure on firms’ innovative performance” Ortega-Argilés, R.; Moreno, R.
- WP 2009/08 “¿Por qué se privatizan servicios en los municipios (pequeños)? Evidencia empírica sobre residuos sólidos y agua” Bel, G.; Fageda, X.; Mur, M.
- WP 2009/07 “Empirical analysis of solid management waste costs: Some evidence from Galicia, Spain” Bel, G.; Fageda, X.
- WP 2009/06 “Intercontinental flights from European Airports: Towards hub concentration or not?” Bel, G.; Fageda, X.
- WP 2009/05 “Factors explaining urban transport systems in large European cities: A cross-sectional approach” Albalate, D.; Bel, G.

- WP 2009/04 “Regional economic growth and human capital: the role of overeducation” Ramos, R.; Suriñach, J.; Artís, M.
- WP 2009/03 “Regional heterogeneity in wage distributions. Evidence from Spain” Motellón, E.; López-Bazo, E.; El-Attar, M.
- WP 2009/02 “Modelling the disability severity score in motor insurance claims: an application to the Spanish case” Santolino, M.; Boucher, J.P.
- WP 2009/01 “Quality in work and aggregate productivity” Royuela, V.; Suriñach, J.
- WP 2008/16 “Intermunicipal cooperation and privatization of solid waste services among small municipalities in Spain” Bel, G.; Mur, M.
- WP 2008/15 “Similar problems, different solutions: Comparing refuse collection in the Netherlands and Spain” Bel, G.; Dijkgraaf, E.; Fageda, X.; Gradus, R.
- WP 2008/14 “Determinants of the decision to appeal against motor bodily injury settlements awarded by Spanish trial courts” Santolino, M.
- WP 2008/13 “Does social capital reinforce technological inputs in the creation of knowledge? Evidence from the Spanish regions” Miguélez, E.; Moreno, R.; Artís, M.
- WP 2008/12 “Testing the FTPL across government tiers” Claeys, P.; Ramos, R.; Suriñach, J.
- WP 2008/11 “Internet Banking in Europe: a comparative analysis” Arnaboldi, F.; Claeys, P.
- WP 2008/10 “Fiscal policy and interest rates: the role of financial and economic integration” Claeys, P.; Moreno, R.; Suriñach, J.
- WP 2008/09 “Health of Immigrants in European countries” Solé-Auró, A.; M.Crimmins, E.
- WP 2008/08 “The Role of Firm Size in Training Provision Decisions: evidence from Spain” Castany, L.
- WP 2008/07 “Forecasting the maximum compensation offer in the automobile BI claims negotiation process” Ayuso, M.; Santolino, M.
- WP 2008/06 “Prediction of individual automobile RBNS claim reserves in the context of Solvency II” Ayuso, M.; Santolino, M.
- WP 2008/05 “Panel Data Stochastic Convergence Analysis of the Mexican Regions” Carrion-i-Silvestre, J.L.; German-Soto, V.
- WP 2008/04 “Local privatization, intermunicipal cooperation, transaction costs and political interests: Evidence from Spain” Bel, G.; Fageda, X.
- WP 2008/03 “Choosing hybrid organizations for local services delivery: An empirical analysis of partial privatization” Bel, G.; Fageda, X.
- WP 2008/02 “Motorways, tolls and road safety. Evidence from European Panel Data” Albalade, D.; Bel, G.
- WP 2008/01 “Shaping urban traffic patterns through congestion charging: What factors drive success or failure?” Albalade, D.; Bel, G.
- WP 2007/19 “La distribución regional de la temporalidad en España. Análisis de sus determinantes” Motellón, E.
- WP 2007/18 “Regional returns to physical capital: are they conditioned by educational attainment?” López-Bazo, E.; Moreno, R.
- WP 2007/17 “Does human capital stimulate investment in physical capital? evidence from a cost system framework” López-Bazo, E.; Moreno, R.
- WP 2007/16 “Do innovation and human capital explain the productivity gap between small and large firms?” Castany, L.; López-Bazo, E.; Moreno, R.

- WP 2007/15 “Estimating the effects of fiscal policy under the budget constraint” Claeys, P.
- WP 2007/14 “Fiscal sustainability across government tiers: an assessment of soft budget constraints” Claeys, P.; Ramos, R.; Suriñach, J.
- WP 2007/13 “The institutional vs. the academic definition of the quality of work life. What is the focus of the European Commission?” Royuela, V.; López-Tamayo, J.; Suriñach, J.
- WP 2007/12 “Cambios en la distribución salarial en España, 1995-2002. Efectos a través del tipo de contrato” Motellón, E.; López-Bazo, E.; El-Attar, M.
- WP 2007/11 “EU-15 sovereign governments’ cost of borrowing after seven years of monetary union” Gómez-Puig, M.
- WP 2007/10 “Another Look at the Null of Stationary Real Exchange Rates: Panel Data with Structural Breaks and Cross-section Dependence” Syed A. Basher; Carrion-i-Silvestre, J.L.
- WP 2007/09 “Multicointegration, polynomial cointegration and I(2) cointegration with structural breaks. An application to the sustainability of the US external deficit” Berenguer-Rico, V.; Carrion-i-Silvestre, J.L.
- WP 2007/08 “Has concentration evolved similarly in manufacturing and services? A sensitivity analysis” Ruiz-Valenzuela, J.; Moreno-Serrano, R.; Vaya-Valcarce, E.
- WP 2007/07 “Defining housing market areas using commuting and migration algorithms. Catalonia (Spain) as an applied case study” Royuela, C.; Vargas, M.
- WP 2007/06 “Regulating Concessions of Toll Motorways, An Empirical Study on Fixed vs. Variable Term Contracts” Albalade, D.; Bel, G.
- WP 2007/05 “Decomposing differences in total factor productivity across firm size” Castany, L.; Lopez-Bazo, E.; Moreno, R.
- WP 2007/04 “Privatization and Regulation of Toll Motorways in Europe” Albalade, D.; Bel, G.; Fageda, X.
- WP 2007/03 “Is the influence of quality of life on urban growth non-stationary in space? A case study of Barcelona” Royuela, V.; Moreno, R.; Vayá, E.
- WP 2007/02 “Sustainability of EU fiscal policies. A panel test” Claeys, P.
- WP 2007/01 “Research networks and scientific production in Economics: The recent Spanish experience” Duque, J.C.; Ramos, R.; Royuela, V.
- WP 2006/10 “Term structure of interest rate. European financial integration” Fontanals-Albiol, H.; Ruiz-Dotras, E.; Bolancé-Losilla, C.
- WP 2006/09 “Patrones de publicación internacional (SSCI) de los autores afiliados a universidades españolas, en el ámbito económico-empresarial (1994-2004)” Suriñach, J.; Duque, J.C.; Royuela, V.
- WP 2006/08 “Supervised regionalization methods: A survey” Duque, J.C.; Ramos, R.; Suriñach, J.
- WP 2006/07 “Against the mainstream: Nazi privatization in 1930s Germany” Bel, G.
- WP 2006/06 “Economía Urbana y Calidad de Vida. Una revisión del estado del conocimiento en España” Royuela, V.; Lambiri, D.; Biagi, B.
- WP 2006/05 “Calculation of the variance in surveys of the economic climate” Alcañiz, M.; Costa, A.; Guillén, M.; Luna, C.; Rovira, C.
- WP 2006/04 “Time-varying effects when analysing customer lifetime duration: application to the insurance market” Guillen, M.; Nielsen, J.P.; Scheike, T.; Perez-Marin, A.M.
- WP 2006/03 “Lowering blood alcohol content levels to save lives the European experience” Albalade, D.



- WP 2006/02 “An analysis of the determinants in economics and business publications by spanish universities between 1994 and 2004” Ramos, R.; Royuela, V.; Suriñach, J.
- WP 2006/01 “Job losses, outsourcing and relocation: empirical evidence using microdata” Artís, M.; Ramos, R.; Suriñach, J.



Institut de Recerca en Economia Aplicada Regional i Pública  
*Research Institute of Applied Economics*

**Universitat de Barcelona**

Av. Diagonal, 690 • 08034 Barcelona

**WEBSITE:** [www.ub.edu/irea/](http://www.ub.edu/irea/) • **CONTACT:** [irea@ub.edu](mailto:irea@ub.edu)