



HAL
open science

RESEARCH NETWORKS AND INVENTORS' MOBILITY AS DRIVERS OF INNOVATION: EVIDENCE FROM EUROPE

Ernest Miguelez, Rosina Moreno

► **To cite this version:**

Ernest Miguelez, Rosina Moreno. RESEARCH NETWORKS AND INVENTORS' MOBILITY AS DRIVERS OF INNOVATION: EVIDENCE FROM EUROPE. *Regional Studies*, 2011, pp.1. 10.1080/00343404.2011.618803 . hal-00749641

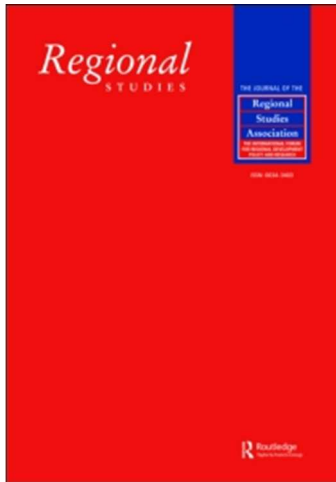
HAL Id: hal-00749641

<https://hal.science/hal-00749641>

Submitted on 8 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



RESEARCH NETWORKS AND INVENTORS' MOBILITY AS DRIVERS OF INNOVATION: EVIDENCE FROM EUROPE

| | |
|------------------|--|
| Journal: | <i>Regional Studies</i> |
| Manuscript ID: | CRES-2010-0038.R2 |
| Manuscript Type: | Special Issue Paper |
| JEL codes: | J61 - Geographic Labor Mobility Immigrant Workers < J6 - Mobility, Unemployment, and Vacancies < J - Labor and Demographic Economics, O31 - Innovation and Invention: Processes and Incentives < O3 - Technological Change Research and Development < O - Economic Development, Technological Change, and Growth, O33 - Technological Change: Choices and Consequences Diffusion Processes < O3 - Technological Change Research and Development < O - Economic Development, Technological Change, and Growth, R1 - General Regional Economics < R - Urban, Rural, and Regional Economics |
| Keywords: | inventors' mobility, networks of co-inventors, knowledge production function, spatial econometrics, European regions |
| | |

SCHOLARONE™
Manuscripts

1
2
3 RESEARCH NETWORKS AND INVENTORS' MOBILITY AS
4
5
6 DRIVERS OF INNOVATION: EVIDENCE FROM EUROPE
7
8
9
10

11 Ernest Miguélez. AQR-IREA. Department of Econometrics, Statistics and Spanish
12 Economy. University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain. E-mail:
13 emiguel@ub.edu.
14
15
16
17

18
19
20 Rosina Moreno. AQR-IREA. Department of Econometrics, Statistics and Spanish
21 Economy. University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain. E-mail:
22 rmoreno@ub.edu.
23
24
25
26

27 (Received January 2011; in revised form July 2011)
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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, a knowledge production function framework at the regional level is used. The empirical approach presented here 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 (SHAC) techniques. The results suggest the existence of a robust positive correlation between intra-regional labour mobility and regional innovation, whilst the relationship with networks is less clear.

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

JEL: J61, O31, O33, R1

1. Introduction

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

1
2
3 features of the local labour market and regional innovative capability is investigated.
4
5 Specifically, the empirical analysis is based on a regional knowledge production function
6
7 (KPF hereafter) framework, which is used to model the aforementioned relationship for a
8
9 large sample of European NUTS-2 regions.
10

11
12
13
14 Drawing on patent data and computerized algorithms to identify individual inventors, a
15
16 large dataset of individuals containing information such as their personal address(es), their
17
18 patenting history, the owners of their patents (i.e. a firm, a university or other public
19
20 institution, or the inventors themselves), the co-authors of their patents, and so on, is
21
22 created. Given the lack of data on single, individual inventors, very few studies have
23
24 examined the influence of these features on regional innovation.
25
26

27
28
29 Among the questions addressed in this study are the following: What is the contribution
30
31 of inventor networking and inventors' labour mobility to the regional intensity of
32
33 patenting? Do cross-regional mobility and cross-regional networking play an important
34
35 role? What impact is attributable to mobility and networking once spatial interactions
36
37 have been controlled for?
38
39

40
41
42 Our choice of the term knowledge "flows" instead of, for instance, knowledge "spillovers"
43
44 or "externalities" is not accidental and requires an explanation before proceed. As
45
46 BRESCHI and LISSONI (2009) point out, knowledge may flow locally, but not necessarily
47
48 as an externality. When knowledge is transferred following the hiring of an employee
49
50 from another firm, or by means of a cooperation agreement, market mechanisms are
51
52 operating and no externalities or spillovers can be said to exist. Clearly, these flows may
53
54 include pure externalities if all the knowledge transferred is not compensated accordingly,
55
56 but in all instances it will only constitute a marginal part. With these thoughts in mind, the
57
58
59
60

1
2
3 initial intention of the present empirical approach was to disentangle the effects of market
4 and non-market knowledge flows on innovation in a multivariate estimation framework.
5
6
7 However, it is very difficult to do this empirically, first because the mechanisms defined as
8
9 market-based may bring pure externalities with them, and second because the ones
10
11 acknowledged as pure spillovers, e.g., agglomeration externalities, may be partially paid
12
13 off. So this is an important theoretical discussion that cannot be fully addressed in an
14
15 empirical framework like the present one.¹
16
17
18
19

20
21 The econometric approach takes full account of the spatial structure of the data by
22
23 applying spatial econometric techniques. Unlike previous studies of innovation, however,
24
25 the spatial lag model is not used to the exclusion of the error model or vice-versa, but
26
27 rather a rich spatial specification is allowed for. To achieve this, first the spatial lag model
28
29 is estimated and then the non-parametric heteroskedasticity and autocorrelation
30
31 consistent (HAC) estimator of the variance-covariance (V-C) matrix in a spatial framework
32
33 (following KELEJIAN and PRUCHA, 2007) is applied is performed. To the best of our
34
35 knowledge, no previous studies have estimated a knowledge production function using
36
37 both a spatial lag model and the robust spatial HAC estimation of the V-C matrix.
38
39
40
41

42
43 The rest of the paper is organized as follows: section 2 reviews the literature on inventor
44
45 networking and mobility, and their relationship with knowledge diffusion and creation;
46
47 section 3 describes the model and the hypotheses proposed; section 4 presents the data;
48
49 whilst section 5 includes the results. Finally, section 6 presents the conclusions and
50
51 identifies certain limitations in the approach.
52
53
54
55
56
57

58 2. Background

59
60

1
2
3
4
5 The literature on collaborative research networks, and their impact on knowledge
6 diffusion and innovation, has expanded greatly in recent years.² This is particularly true in
7 the case of networks of co-inventors, thanks to the availability of relevant data (that is, co-
8 patent data). Part of this literature has been devoted to explaining the determinants of
9 these collaborative patterns (HOEKMAN et al. 2009; MAGGIONI and UBERTI, 2008), while
10 another important strand has focused on networks as mechanisms for inter-regional R&D
11 spillovers (KROLL, 2009; PONDS et al., 2007, 2010), and, in particular, networks as the
12 means by which knowledge diffuses between individuals and across firms (BRESCHI and
13 LISSONI, 2004, 2006; 2009; GOMES-CASSERES et al., 2006; SINGH, 2005).

14
15
16
17
18
19
20
21
22
23
24
25
26
27 SINGH (2005) finds strong evidence in the US that the existence of interpersonal ties in
28 the form of co-patents increases the probability of knowledge flows, as measured by
29 patent citations. Indeed, he hints that these ties are important in determining patterns of
30 intra-regional and intra-firm knowledge flows. SINGH claims that geography matters
31 especially because interpersonal networks tend to be regional in nature (Op. cit.). Similar
32 results are found by BRESCHI and LISSONI (2004) for Italy. Recent findings by BRESCHI
33 and LISSONI (2009), using patent applications to the European Patent Office (EPO) in
34 certain technological fields (biotechnology, organic chemistry, and pharmaceuticals),
35 suggest that networking activity across firms (and locations) is to a large extent responsible
36 for the localisation of knowledge flows, indicating that the residual effect of non-market
37 externalities is not as great as was previously believed.

38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54 All these studies stress the importance of networks as mechanisms of knowledge
55 transmission, and hence of knowledge creation. Co-location and shared space are
56 reported to be neither necessary nor sufficient for knowledge flows; rather it appears to be
57
58
59
60

1
2
3 social distance, or the quality and quantity of interactions between those individuals most
4
5 closely involved with innovative activities, that is critical for the effective diffusion of
6
7 knowledge (BOSCHMA, 2005). If this is the case, the features of the inventors' network
8
9 structure at any given location will play a significant role in regional innovation outcomes.
10
11 A number of empirical analyses have recently been conducted in a knowledge production
12
13 function framework by BETTENCOURT et al. (2007a,b), FLEMING et al. (2007), and LOBO
14
15 and STRUMSKY (2008) for the case of US metropolitan statistical areas (MSAs). Broadly
16
17 speaking, various features of the inventors' network are found to have different and
18
19 significant effects on creativity. While the scale and extent of networks are reported to
20
21 have a positive impact, the density of these networks seems to have a negative effect,
22
23 which has been attributed to the existence of superfluous information (BETTENCOURT et
24
25 al., 2007a,b; LOBO and STRUMSKY, 2008). However, the importance of inventors'
26
27 networks is not equally significant in all the analyses conducted (FLEMING et al., 2007),
28
29 and further research in this field is still required. At present, few empirical studies of
30
31 inventors' networks have been performed and the evidence of their benefits remains
32
33 scarce.
34
35
36
37
38
39

40
41 The present analysis is closely related to these recent studies, since it also uses a KPF to
42
43 assess the effect of networks and their particular structure on aggregate innovation
44
45 outcomes. However, here the production function is applied to European regions as
46
47 opposed to US MSAs.³ Moreover, in contrast to the earlier studies, spatial relationships
48
49 across regions are taken fully into account by applying spatial econometric techniques.
50
51 Yet, no significantly different results are expected a priori vis-à-vis those previously
52
53 reported for the US.
54
55
56
57
58
59
60

1
2
3 Similarly, earlier studies have examined how the labour mobility of inventors acts as a key
4
5 mechanism in the diffusion of knowledge (ALMEIDA and KOGUT, 1999; ROSENKOPF and
6
7 ALMEIDA, 2003; SAXENIAN, 1994). However, this line of research has been less prolific.⁴
8

9
10 One strand of this literature has shown the relationship between mobility and the flow of
11
12 knowledge as measured by patent citations, as well as the knowledge gain by a firm hiring
13
14 an inventor from another firm. For example, in a pioneering study, ALMEIDA and
15
16 KOGUT (1999) show that inter-firm mobility of patent holders in the US semiconductor
17
18 industry influences the local transfer of knowledge across firms. Similar findings are
19
20 reported in BRESCHI and LISSONI (2009)'s study mentioned above for US inventors in
21
22 selected technological fields making patent applications to the EPO. In a similar vein,
23
24 AGRAWAL et al. (2006) stress that when inventors leave their workplace they maintain
25
26 interpersonal ties with their former colleagues, who may well later cite their work. In
27
28 addition, several studies (CRESPI et al., 2007; CORREDORIA and ROSENKOPF, 2006; KIM
29
30 et al. 2006; SINGH and AGRAWAL, 2011; SONG et al., 2003) have stressed the role of
31
32 mobility in increasing the hiring firm's use of a hired inventor's prior knowledge.
33
34
35
36
37

38 Parallel to these studies, another line of research has studied mobility by focusing
39
40 attention on inventors' performance itself. For instance, mobility-productivity
41
42 relationships have been studied by HOISL (2007, 2009) for German, by LENZI (2009) for
43
44 Italian, and by SHALEM and TRAJTENBERG (2008) for Israeli inventors. Broadly speaking,
45
46 it has been shown that mobility may in fact enhance productivity (HOISL, 2007), although
47
48 results in this direction are not as robust as one would expect (SHALEM and
49
50 TRAJTENBERG, 2008), a fact that these authors attributed to what they identified as the
51
52 short-term costs of mobility.
53
54
55
56
57
58
59
60

1
2
3 If an individual's innovative output increases on moving, and if the new host firms acquire
4 more knowledge and are more efficient in their innovative activities, the innovative
5 capacity of a region as a whole should increase as the degree of inventors' mobility rises
6 across firms within a region. To the best of our knowledge, there have been few empirical
7 attempts at quantifying the effect of this feature of the local labour market for inventors -
8 in other words, the degree of their job-to-job mobility - on regional innovation outcomes.
9
10 As such, this constitutes one of the main contributions of this paper.
11
12
13
14
15
16
17
18
19
20
21
22

23 3. Model and hypotheses

24 25 26 27 Baseline specification

28
29
30
31 In order to meet the goals identified above in the introduction, the knowledge production
32 function framework at the regional level is used. First used in the seminal studies of
33 GRILICHES (1979) and HAUSMAN et al. (1984) at the firm level, this framework was
34 subsequently extended by JAFFE (1986, 1989) to the regional level. Since then, regional
35 KPF has been the standard tool for assessing the influence of regional innovation efforts
36 on regional innovation output. However, many empirical exercises have sought to test this
37 relationship by simultaneously considering a number of regional structural characteristics,
38 which would also tend to have a positive impact on regional innovation output. The KPF
39 has been widely estimated for the case of the US (ACS et al., 2002; ANSELIN et al., 1997)
40 and Europe (ACKÇOMACK and TER WEEL, 2009; BOTTAZZI and PERI, 2003; DE
41 DOMINICIS et al, 2007; MORENO et al., 2005a).
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

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

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, the innovation output is allowed to depend on an R&D productivity index³ for that region, A . For simplicity, it is assumed that the KPF follows a Cobb-Douglas functional form:

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

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 the estimations, the regional observations need to be comparable in size: equation 2 is pre-multiplied by a factor of $1/N$, where N is the total population of each region.

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

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 kinds 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

1
2
3 actually account for how innovation is generated at the regional level (RODRIGUEZ-POSE
4 and CRESCENZI, 2008). To capture a variety of returns that might affect innovation
5 outcomes, the R&D productivity index is assumed to be a function of a number of (Hicks
6 neutral) control variables. In addition to pure externalities and other controls, this index is
7 assumed to be affected by the level of inventors' labour mobility within a given region,
8 and by the scale and density of its collaborative research networks.
9
10
11
12
13
14
15
16
17

18 The hypotheses to be tested here are:
19
20
21

22 *H1. Collaborative research networks of inventors boost regional innovation capability.*
23
24
25
26

27 Networks of inventors may have a direct influence on innovation. The rationale behind
28 this is that the simple cross-fertilization of previously unconnected ideas will lead to better
29 knowledge outputs (HOEKMAN et al., 2009) and that individuals connected within a
30 collaborative framework are more willing to learn from each other than is the case of
31 isolated inventors. Moreover, collaborative research projects may achieve scale
32 economies and may lower research costs by reducing the duplication of research efforts
33 among the participants in the network. On the other side, indirect effects of networking
34 may be due to knowledge diffusion - and subsequent innovation - and the formation of
35 social capital. Thus, close network links should prove more useful in transferring complex
36 knowledge (COWAND and JONARD, 2004), especially knowledge with a high component
37 of "tacitness" (SINGH, 2005). Additionally, professional relationships of this nature
38 enhance trust and cooperative behaviour between individuals (hence raising the level of
39 social capital) which has been shown to be a further element in innovation and knowledge
40 transmission (AKÇOMAK and TER WEEL, 2009; DE CLERCQ and DAKHLI, 2004;
41 MIGUÉLEZ et al., 2008).
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5 In line with the literature (see Section 2), social network analysis (SNA) tools are
6 employed to investigate empirically the quantitative relationship between inventors'
7 collaborations and levels of inventiveness.⁶ We are particularly interested in measuring
8 two aspects of inventors' networks. The first is the scale of these networks, i.e., whether a
9 greater number of social ties, or just simply more inventors involved in co-innovations, are
10 beneficial for inventive intensity. Here, three distinct variables are used to proxy the scale
11 of the networks in a community. However, we are reluctant to include the three variables
12 at the same time in the regressions because of their high degree of collinearity and so
13 three separate models are estimated. Clearly, a positive effect on creativity is expected.
14
15
16
17
18
19
20
21
22
23
24
25
26

27 The second aspect that concerns us is the strength of the inventors' community ties,
28 measured as the network density. The naïve, expected effect of density on innovation is
29 positive. However, we should bear in mind GRANOVETTER's (1985) warning that overly
30 strong interpersonal ties may well hamper innovation because, at some point, the
31 information flowing across those ties becomes redundant. As shown in Section 5 below,
32 the findings of the empirical analysis of network density are in agreement with those
33 encountered for the case of the US.
34
35
36
37
38
39
40
41
42
43
44

45 *H2. The mobility of inventors within the local labour market of a region enhances*
46 *innovative intensity.*
47
48
49
50

51 Recent studies have shown that mobile inventors are more productive than non-mobile
52 ones, measured either by patent applications or by citations received to their work (SINGH
53 and AGRAWAL, 2011; HOISL, 2007, 2009; SHALEM and TRAJTENBERG, 2008). Hence,
54 communities of inventors within regions with larger degrees of labour mobility are
55
56
57
58
59
60

1
2
3 expected to be more productive and innovation intensive. On the other hand, mobility
4
5 may also favour knowledge diffusion. As knowledge is embedded in individuals, when
6
7 they move they also take with them the knowledge capital they have accumulated. Their
8
9 movement across firms therefore contributes to knowledge exchange between firms
10
11 (BOSCHMA et al., 2009). Skilled workers take their knowledge with them and share it in a
12
13 new workplace with their new employer and colleagues. In return, they acquire new
14
15 knowledge from their new colleagues, establish new links and social networks for future
16
17 collaborations based on trust and, in general, promote new combinations of knowledge
18
19 (LAUDEL, 2003; TRIPPL and MAIER 2010).
20
21

22
23
24 Given these two hypotheses, the R&D productivity index is a function of the following
25
26 variables
27
28

$$29 \quad A_i = g(A_C, MOB_i, NET_i, DENS_i, Z_i, X_i), \quad (4)$$

30
31
32 where MOB is the measure of mobility, NET is the scale of the networks in a given
33
34 region, $DENS$ is a measure of the density of regional networking, Z is a set of variables
35
36 controlling for the existence of externalities and other knowledge flow mechanisms
37
38 derived from agglomeration economies, specialization, urbanization, and market access;
39
40 and X is a set of technical controls to capture differences in economic structures across
41
42 regions. Moreover, A_C is a country-specific effect capturing both institutional differences
43
44 across countries affecting innovation and the existence of so-called national systems of
45
46 innovation. Thus, equation 3 becomes
47
48
49
50
51
52
53
54

$$55 \quad y_i = C \cdot [g(A_C, MOB_i, NET_i, DENS_i, Z_i, X_i)]^\delta \cdot k_i^\beta \cdot l_i^\alpha. \quad (5)$$

The R&D productivity index is modelled 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 in the resulting expression gives the following empirical model, to which an (a priori) random error term is now added, ε_i :

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

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 a cross-section regression is going to be estimated, we should be aware that all the explanatory variables pre-date the period of analysis of the dependent variable and can be assumed, in principle, to be exogenous.

Spatial structure of innovation

1
2
3 It is acknowledged that innovation and knowledge variables are, by and large, spatially
4 correlated, irrespective of the level of regional disaggregation, the time span, or the
5 sample of regions analysed (ACS et al., 2002, ANSELIN et al., 1997; AUTANT-BERNARD
6 and LESAGE, 2011; BODE, 2004; MORENO et al., 2005b). Indeed, knowledge tends to
7 flow locally, but there is no reason to assume that it stops flowing because of regional
8 borders (AUDRETSCH and FELDMAN, 2004). Here we are interested in elucidating
9 whether the preliminary findings are robust to the correct specification of these spatial
10 effects. Thus, the following hypothesis is considered:
11
12
13
14
15
16
17
18
19

20
21
22 *H3. The spatial structure of the data also matters.*
23
24

25
26
27 To step in this direction, the R&D productivity index is now assumed to be a function of
28 the innovation intensity of the neighbouring regions of a given region, $\prod_{\substack{j=1 \\ j \neq i}}^N y_j^{\rho w_{ij}}$, where
29
30
31
32
33
34 w_{ij} are bilateral spatial weights, which will depend on the distance between regions i and
35
36
37 j , while ρ is the output elasticity of cross-regional effects of innovation intensity.
38
39
40

$$41 \quad A_i = g(A_C, \prod_{\substack{j=1 \\ j \neq i}}^N y_j^{\rho w_{ij}}, MOB_i, NET_i, DENS_i, Z_i, X_i). \quad (9)$$

42
43
44
45
46
47
48 Expressing this in log form in the multiplicative version of 9 and 3, and making the
49 necessary rearrangements, we obtain the following model for estimation, including once
50 again a random error term, ε_i :
51
52
53
54
55
56
57
58
59
60

$$\begin{aligned}
 \ln PATpc_{it} = & \xi + \rho \sum_{j \neq i} w_{ij} \ln PATpc_{it} + \beta \cdot \ln k_{it-1} + \alpha \cdot \ln l_{it-1} + \omega_1 \cdot \ln MOB_{it-1} + \\
 & + \omega_2 \cdot \ln NET_{it-1} + \omega_3 \cdot \ln DENS_{it-1} + \omega_j \cdot \ln Z_{it-1} + \omega_n \cdot \ln X_{it-1} + \\
 & + \omega_N \cdot \ln DUMMY_COUNTRY_EFFECT + \varepsilon_i
 \end{aligned} \quad (10)$$

Thus, equation (10) suggests the estimation of the spatial lag model - as opposed to the spatial error model (see ANSELIN, 1988). Indeed, this model implies an economically meaningful interpretation, i.e., spillovers across regions captured in the spatial lag of the dependent variable, whereas the spatial error autocorrelation can be attributed to noise (ANSELIN, 1988) or even to a misspecification of the model (ANDERSSON and GRÅSJÖ, 2009). However, checks for any remaining spatial correlation in the residuals following the estimation of this model will be presented.

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, these 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 need to estimate the spatial lag and error model simultaneously will make the use of the technique cumbersome, if not impossible (FINGLETON and LE GALLO, 2008). In this setting, spatial two stage least squares (S2SLS) procedures are preferred, where the instruments are the spatial lags of the remaining explanatory variables.⁷

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 the present paper a

1
2
3 geographical weights matrix is expressly specified, although we are aware of the criticism
4
5 that these matrices have received from a number of scholars who claim they are
6
7 oversimplified (BOSCHMA, 2005). Nonetheless, as is shown in the next subsection, the
8
9 interest is precisely in identifying what part is attributable to geographical space when
10
11 other more meaningful variables describing interactions across regions are considered.
12
13 However, instead of simply using contiguity matrices, the concept of k-nearest neighbours
14
15 is used. Thus, we measure the distance to a fixed number of neighbours, k, using a matrix
16
17 where k=5, because both the mean and the median number of neighbours for the
18
19 contiguity matrix is located around 5.
20
21
22
23
24

25 Networks and mobility: Cross-regional linkages
26
27
28

29 Continuing with the idea of exploiting specific relationships across regions, explicit inter-
30
31 linkages are now taken into account. If they are relevant for regional innovation, their
32
33 omission will lead to biased and inconsistent estimates.
34
35
36
37

38 *H4. Cross-regional collaborations and inter-regional mobility enhance creativity. The*
39
40 *more inventors collaborate with fellow inventors outside the region, the greater are the*
41
42 *returns on innovation. Similarly, the greater the number of inventors moving into a given*
43
44 *region, the greater is the patenting activity of that region.*
45
46
47
48

49 To check this hypothesis, two additional variables are now included. Thus, the last
50
51 specification includes the variable OUT_COLL, which compares the number of co-
52
53 inventors outside the region with the total number of inventors within a region. They also
54
55 include the variable Inward Migration Rate (IMR), which is calculated counting the
56
57 number of inventors moving into the region compared to the total number of inventors
58
59
60

1
2
3 identified in that region. Other things being equal, a positive and significant effect for both
4
5 variables, for the same reason that a positive effect was expected for the scale of networks
6
7 and labour mobility within the region, is expected. This cross-regional knowledge is likely
8
9 to be critical for innovation because imported knowledge is even more relevant than that
10
11 already held locally.
12

13 14 15 16 17 18 4. Data 19

20
21
22 Innovation intensity is measured by patent applications per capita (PAT_{pc}), a variable
23
24 widely used in the literature to proxy innovation outcomes. As is well known, this proxy
25
26 presents serious caveats since not all inventions are patented, nor do they all have the
27
28 same economic impact, as they are not all commercially exploitable (GRILICHES, 1991).
29
30 In spite of these shortcomings, patent data have proved useful for proxying inventiveness
31
32 as they present minimal standards of novelty, originality and potential profits, and as such
33
34 are a good proxy for economically profitable ideas (BOTTAZZI and PERI, 2003). Patent
35
36 data come from the REGPAT database (January 2009 edition), which is the result of the
37
38 OECD's regionalisation of the PATSTAT dataset (all EPO patents). Since these data are
39
40 prone to exhibit lumpiness from year to year, we have averaged out patent figures from
41
42 2001 to 2003, thereby mitigating the effects of annual fluctuations in this variable,
43
44 especially in the less populated areas. The same procedure is repeated for the explanatory
45
46 variables, for the period 1996-2000.⁸
47
48
49
50

51
52
53 R&D expenditure data are drawn from Eurostat and are used to construct the R&D stock
54
55 proxy (STOCK_R&D_{pc}). The accumulation of R&D is described as
56
57 $\Delta R \& D = R \& D_t - \delta R \& D_{t-1}$ - see PERI (2005) - where δ is an arbitrary depreciation
58
59
60

1
2
3 rate set at 5%. Other depreciation rates, however, do not affect the results. The initial
4
5 stock of R&D has been calculated using the perpetual inventory method,
6
7 $R \& D_0 = R \& D_{t-n} / (g + \delta)$, where $R \& D_{t-n}$ is the first year for which the data were
8
9 available, and g is the geometric average annual growth rate of R&D expenditure
10
11 between 1996 and 2000. R&D employment has been proxied using Human Resources in
12
13 Science and Technology (HRST) data extracted from Eurostat. Ideally, direct data on
14
15 R&D employment, also available from Eurostat, should be used, but the number of
16
17 missing figures is considerable and so we opted not to use this dataset.
18
19

20
21
22
23 The data for constructing the mobility and network variables are also taken from the
24
25 REGPAT database. In spite of the vast amount of information contained in patent
26
27 documents, a single ID for each inventor and anyone else is missing. However, in order
28
29 to draw the mobility and networking history of inventors, it is necessary to identify them
30
31 individually by name and surname, as well as via the other useful details contained in the
32
33 patent document. The method chosen for identifying the inventors is therefore of the
34
35 utmost importance in studies of this nature. Here, then, the methodology proposed by
36
37 MIGUÉLEZ and GÓMEZ-MIGUÉLEZ (2010) is followed, who, in line with a growing
38
39 number of researchers in the field, suggest several algorithms for singling out individual
40
41 inventors using patent documents. In the present study, this procedure has been used for
42
43 a subsample of inventors whose patent applications have been made from one of 17
44
45 countries (see Appendix 1) between 1993 and 2002. Additionally, and unlike the
46
47 dependent variable, we only apply the algorithms to EPO patents also filed under the
48
49 Patent Cooperation Treaty (PCT). This condition was adopted for two main reasons. The
50
51 first is that patents filed during an international phase (the PCT) can be assumed to be
52
53 more technologically and economically exploitable, since applying for EPO patents under
54
55 PCT procedures is more expensive and time-consuming; therefore, the knowledge
56
57
58
59
60

1
2
3 embedded in these patents is presumably of greater worth for innovative activities. The
4
5 second reason is related to time constraints, because the chosen methodology still
6
7 requires a considerable amount of manual work to ensure minimum levels of reliability.
8
9 Eventually, however, the procedure provides us with a sample of 646,678 records, from
10
11 which 285,831 individual inventors are identified.
12

13
14
15
16 Once each inventor has been assigned an individual identification, mobility and network
17
18 data can be calculated for each region. A “mobile” inventor is broadly defined as an
19
20 individual who moves across different organisations offering his/her services (BRESCHI
21
22 and LISSONI, 2009). Therefore, mobility can refer either to labour mobility understood in
23
24 its strictest sense (an employee leaving a firm to take up a position in a new one), or to
25
26 that demonstrated by consultants, freelance workers, university inventors, and the like.
27
28 We assume that both constitute sources of knowledge flows to the extent that in the two
29
30 instances knowledge is transferred from former employers or customers to new ones.
31
32 Mobility is then proxied as the average number of assignees⁹ for each inventor identified
33
34 per region. To make this calculation, the recent study by THOMA et al. (2009a,b) and the
35
36 OECD Harmonised Applicants’ Names (HAN) database (October, 2009), aimed at
37
38 identifying single applicants in patent documents, is used. Additionally, all patents with
39
40 more than one applicant, or assignee, are discarded, as we were unable to discriminate
41
42 labour relationships between applicants and inventors in a multi-applicant patent.
43
44 However, from the initial sample of 646,678 records, only 59,687 were discarded.
45
46
47
48
49
50

51
52 The design of the network variables is based on the theory of SNA. The inventors form
53
54 the nodes in the network, and these are grouped via edges or ties (in this instance, co-
55
56 patents) into different components. A component, therefore, is formed by all the
57
58 inventors in the network who are linked in such a way that a continuous path via the edges
59
60

1
2
3 can be traced across them. In this way, even if a given pair of inventors has not co-
4 patented an invention but they share a co-invention with a third inventor, they will form
5 part of the component: i.e., components may comprise both direct and indirect links.
6
7 The largest component is the largest set of inventors that can be linked through direct and
8 indirect collaboration ties.
9
10
11
12

13
14
15
16 The scale of the network structure in each region – each regional community of inventors
17 – is proxied using three different variables. Among them, two measures of connectivity of
18 the network are included. Connectivity is measured counting the number of edges (ties)
19 between inventors within a given region. Thus, both the absolute number of ties
20 (CONN_ABS) and its relative measure (CONN_INV – the absolute number of ties with
21 respect to the number of inventors within each region) are considered separately. As a
22 third measure, the size of the largest component (SIZE_LC – the number of inventors
23 that belong to the largest component with respect to the total number of inventors, in
24 each region) is included as well. A positive and significant coefficient is expected for both
25 the mobility and the networking variables.
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 The strength of these ties is proxied by the network density, which is the number of ties
41 between inventors within the region divided by the possible number of ties within that
42 region. Formally,
43
44
45
46
47
48

$$49 \quad DENS_i = \frac{T_i}{Q_i(Q_i - 1)/2}, \quad (11)$$

50
51
52
53
54
55
56
57
58
59
60

1
2
3 where T_i stands for the number of edges (ties) within a given region, and Q_i is the total
4
5 number of inventors within that region. As stressed earlier, the expected effect (be it
6
7 positive or negative) of innovation density is not so clear a priori.
8
9

10
11 In the specification, up to four variables are included to capture the effect of other
12
13 knowledge flow mechanisms (Z). Thus, gross value added per capita (capturing market
14
15 access, general economic conditions, and wealth - GVApc), the number of technological
16
17 firms to the total number of manufacturing establishments in that region (capturing
18
19 agglomeration of firms engaged in patenting activity - TECH_FIRMS), population
20
21 density (capturing agglomeration economies and urbanization economies - inter-industry
22
23 economies - POP_DENS), and a technological specialization index (capturing
24
25 specialization economies - intra-industry economies - constructed using the eight
26
27 technological classes of the IPC¹⁰ classification¹¹ - SPIN) are considered in the regressions.
28
29
30
31

32
33 Likewise, a set of four controls (X) capture differences in the technological and
34
35 economic structures across regions: the shares of manufacturing employment in each
36
37 region (MAN), and the shares of biotechnology (BIOTECH), organic chemistry
38
39 (CHEM), and pharmaceuticals (PHARMA) in their patenting activity, according to the
40
41 IPC classification; since these three sectors tend to be more research intensive, controls
42
43 for their presence in each region are necessary.¹² A summary of the variable descriptions
44
45 and the data sources can be found in Appendix 2.¹³
46
47
48
49

50
51
52 The empirical analysis is conducted for the NUTS 2 regions of 17 Western European
53
54 countries (EU-15 plus Norway and Switzerland, see Appendix 1). However, due to data
55
56
57
58
59
60

1
2
3 constraints NUTS 1 regions have also been considered for the UK, and NUTS 0 for
4
5 Denmark and Switzerland.¹⁴
6
7
8
9

10 11 5. Results 12

13
14
15
16 Results on the role of research networks and labour mobility of inventors
17
18
19

20
21 Table 1 presents the results of the OLS estimations with robust standard errors to control
22 for heteroskedasticity of unknown form. In columns (i), (ii), and (iii) the estimations are
23 presented without country dummies. It should be borne in mind that three specifications
24 using alternative proxies for the network scale have been estimated. In principle, the
25 coefficients can be interpreted as elasticities, since the variables in the regression are
26 expressed in natural logarithmic form: the proportional increase in patenting activity in
27 response to a 1% increase in a given explanatory variable. In this way, the interpretation of
28 R&D stock and human capital coefficients is straightforward. In the case of the remaining
29 coefficients, some doubts must remain regarding their direct interpretation. Since the
30 functional relationship between these variables and their response is decided in an ad-hoc
31 fashion, their sign and significance can be fairly informative, although any interpretation of
32 their magnitude should be treated with caution (BODE, 2004).
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Some specific results are worth highlighting. The R&D stock and human capital
50 parameters are, in general, strongly significant; they have the expected sign and present
51 similar values to those reported elsewhere in the literature. The control variables tend to
52 be significant and with the expected sign, although there are a number of exceptions. The
53 specialisation index is only significant (and positive) when country dummies are included,
54
55
56
57
58
59
60

1
2
3 and the same is true in the subsequent tables. The share of patents in organic chemistry
4
5 has a positive impact on innovation, whilst, surprisingly, biotechnology and
6
7 pharmaceutical shares correlate negatively (and in the latter case, the correlation is always
8
9 significant). We believe that manufacturing employment (strongly positive and significant)
10
11 must capture part of the effect of these three sectors on innovation. Strikingly, population
12
13 density is significant, but contrary to our expectations, it is negative. This can be
14
15 interpreted in several ways: first, the positive effects of market potential may already be
16
17 taken into account with the inclusion of GVA per capita, which is always significant and
18
19 positive. Moreover, it might be that overly dense areas suffer several costs related to the
20
21 congestion effects of agglomeration (negative externalities). Alternatively, the result may
22
23 also reflect the fact that the regions in the sample are particularly heterogeneous. Finally,
24
25 the significant yet negative effect of population density on innovation may point to the
26
27 absence of urbanisation (inter-industry) economies, at the expense of specialisation
28
29 economies.¹⁵
30
31
32
33
34
35

36 The focus variables of this study are also significant. Mobility, for example, is always
37
38 significant at 1%, presenting parameters around 0.26 and 0.29, whilst the relationship
39
40 between the scale of the networks and innovation is always positive and strongly
41
42 significant, except in the case of the size of the largest component variable. Thus,
43
44 hypotheses 1 and 2 of the theoretical section are confirmed. In addition, network density
45
46 shows a negative impact on innovation intensity, which supports GRANOVETTER's (1985)
47
48 arguments about weak ties and innovation. Encouragingly, the results are broadly similar
49
50 to the findings reported for the US, although the approach adopted here is slightly
51
52 different.
53
54
55
56
57

58 [Insert Table 1 about here]
59
60

1
2
3
4
5 In short, the empirical analyses undertaken here support the hypotheses concerning the
6 importance of labour mobility and networks in the local labour market for the diffusion
7 and creation of regional innovations. However, several extensions to this initial approach
8 can now be made, in order to relax the initial assumptions concerning the error term of
9 the empirical model.
10
11
12
13
14
15

16
17
18 Results on the spatial structure of the data
19

20
21
22 As can be seen at the end of Table 1, Moran's I tests and Lagrange Multiplier (LM) tests
23 (robust and non-robust) are also calculated. OLS residuals remain spatially correlated
24 after the estimations - even when country dummies are included, although quantitatively
25 their importance is relatively small. Moreover, LM tests seem to point to the need to
26 estimate the spatial lag model, as was suggested in equation 10. Again, however, tests for
27 residual spatial autocorrelation after the estimation of the spatial lag model will be
28 performed.
29
30
31
32
33
34
35
36
37
38
39

40 Thus, Table 2 shows the results of the spatial lag model presented in equation 10, and
41 corresponding to hypothesis 3. As can be seen, the parameter for the spatial lag is positive
42 and strongly significant, suggesting the existence of relevant knowledge flows across
43 regions affecting their respective innovation outcomes and thus confirming the hypothesis.
44 Again, this result is in line with findings reported elsewhere. Comparing these results for
45 the remaining explanatory variables with the previous OLS estimations, it can be seen that
46 both the value of the parameters and their significance remain practically unchanged.
47
48
49
50
51
52
53
54
55
56
57

58 [Insert Table 2 about here]
59
60

1
2
3
4
5 It is shown at the bottom of Table 2 the modified Moran's I test for 2SLS suggested in
6 ANSELIN and KELEJIAN (1997), which points to the existence of a residual correlation
7 (albeit not very strong) in at least three of the estimations (iv, v, and vi). Given these
8 results, we believe that spatial correlation is no longer a problem, although we will
9 confirm this below. To address this issue, recent developments in spatial econometrics by
10 KELEJIAN and PRUCHA (2007) are implemented which, in contrast to earlier work, do not
11 impose a specific functional form of the spatial correlation of the error term¹⁶, i.e. the
12 spatial HAC estimator of the V-C matrix. The rationale behind this technique is drawn
13 from the time-series results; basically it is a non-parametric technique for estimating the V-
14 C matrix using the averages of cross-products of residuals, the range of which is
15 determined by a kernel function. This kernel will determine which pairs of regions, i, j , are
16 included as cross-products. In the present paper, the kernel is formulated as $K(d_{ij}/d)$,
17 with d_{ij} the distance between regions i and j , and d the bandwidth. $K(d_{ij}/d)$ equals 0
18 when $d_{ij} \geq d$. The triangular kernel, $K(d_{ij}/d) = 1 - (d_{ij}/d)$, is used here as the form of
19 the function. The bandwidth used is calculated based on distances to the 10 nearest
20 neighbours. As will be shown in the final section, robustness checks using both different
21 forms of the kernel - Epanechnikov and bisquare - and alternative numbers of nearest
22 neighbours to calculate the bandwidth are performed, but these checks do not alter the
23 results.

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50 Table 3 summarises the results of the estimation of the spatial model with a spatial lag
51 and the spatial heteroskedastic and autocorrelation consistent estimation of the V-C
52 matrix. Encouragingly, the significance of the parameters in Table 2 remains virtually
53 unchanged.

[Insert Table 3 about here]

Results on the existence of cross-regional linkages

Hypothesis 4 is tested by including two additional variables – OUT_COLL and IMR – in all the models. Here, only the estimations with the absolute connectivity (CONN_ABS) variable are shown, though the other models do not change the conclusions greatly. The results (see Table 4 – columns i-vi) corroborate the importance of outside collaborators: even though this variable presents smaller parameters than those presented by other networking variables, it 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. The interpretation of these results is 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., 2010), and thus help to overcome geographic distances (SINGH, 2005). In addition, it could be argued that the actual origin of these inventors – that is, whether or not they come from a more innovative region – also matters.

Additionally, since most of the knowledge flows across regions in the knowledge production process are driven by networks of research collaborations and labour mobility, one would expect the values of the spatial autocorrelation tests to fall significantly when these two variables are included in the regressions. This appears to be

1
2
3 particularly true for the case of the residual spatial correlation tests after the estimation of
4
5 the spatial lag model. If this is indeed the case, a way to model interdependences across
6
7 regions in the production of knowledge has been found.
8
9

10
11 Comparing the spatial residual tests at the bottom of Table 4 with those at the bottom of
12
13 Tables 1 and 2, it is unfortunately impossible to assert categorically that the inclusion of
14
15 economically meaningful linkages across regions has any significant effect on the residuals
16
17 in the previous estimations. In short, a reduction in the OLS spatial residual tests
18
19 (especially those of the spatial LM and RLM) is observed if Table 1 (columns i and iv)
20
21 and Table 4 (columns i and ii) are compared. However, comparing Table 2 (columns i
22
23 and iv) and Table 4 (columns iii and iv), no reduction in the value of the Moran's I is
24
25 observed. Consequently, it seems that the inclusion of these variables does not completely
26
27 eliminate the spatial correlation in the residuals. Obviously, these results may depend on
28
29 the (spatial) weights matrix chosen for the analysis, though several (spatial) alternatives
30
31 have been tried without relevant changes in the results. However, further research on this
32
33 question is required.
34
35
36
37
38
39

40 [Insert Table 4 about here]
41
42
43
44
45
46

47 6. Conclusions, implications and limitations 48 49 50

51 The research conducted here has sought to assess the importance of specific knowledge
52
53 flow mechanisms, namely networks of co-invention and labour mobility, on regional
54
55 innovation, as opposed to the impact of R&D efforts or other mechanisms of knowledge
56
57 creation and diffusion. Within a KPF framework, several hypotheses were suggested and,
58
59
60

1
2
3 although we are unable to confirm them all, a number of interesting conclusions can be
4
5 identified.
6
7

8
9 Strong support for the positive relationship between regional labour market mobility and
10 regional innovation intensity is found. The influence of networks is also fairly important,
11 but the strength of these ties (measured as the network density) was found to have a
12 negative influence on innovation. In line with previous studies, we rely on the
13 explanations proffered by GRANNOVETTER (1985) concerning the importance of weak
14 ties for innovation. This idea is further strengthened with the inclusion in the regressions
15 of a proxy for the number of collaborations with inventors outside the region - hypothesis
16 4, positive and significant - which leads us to conclude that weak, distant ties are indeed
17 more important for innovation.
18
19
20
21
22
23
24
25
26
27
28

29
30
31 Contrary to our expectations, however, the inflow of inventors from other regions does
32 not have a markedly significant impact on innovation outcomes. In principle, these
33 findings can be attributed to the level of regional disaggregation at which the analysis was
34 conducted, although, as SHALEM and TRAJTENBERG (2008) state, the existence of certain
35 short-term costs of a change in location should not be overlooked.
36
37
38
39
40
41
42
43
44

45 However, the results reported here reveal certain tendencies from which a number of
46 policy implications can be directly derived. Research collaborations across firms and
47 regions are pivotal for acquiring external knowledge (see similar claims in SINGH, 2005),
48 and for promoting the creation of new knowledge. The promotion of distant, weak ties
49 embracing as many actors as possible is therefore a plausible and beneficial policy option
50 from a regional perspective. Policy recommendations regarding mobility within the local
51 labour market are not so straightforward. Although mobility seems to be desirable at an
52
53
54
55
56
57
58
59
60

1
2
3 aggregate level, and also at the inventor level, it could be understood as a zero-sum game
4
5 for firms. A policy option might be, therefore, to promote competition for talent at inter-
6
7 regional and, in particular, at international levels.
8
9

10
11 Our approach has several limitations. Below, we note some of the drawbacks to the
12
13 analysis, which should serve as the basis for improved approaches.
14
15

16
17 Patenting activity does not take into account all the research collaboration relationships
18
19 that exist between individuals or all the mobility patterns of those individuals. In so far as
20
21 other variables capturing similar concepts can be assembled, the analysis conducted
22
23 should be repeated. Related to the first of these shortcomings, extending the sample to
24
25 include all EPO patents and not just those that had been filed under the PCT will shed
26
27 more light on the relationships analysed here.
28
29
30
31

32
33 As regards the raw data used, improvements to the suggested algorithms are likely to
34
35 provide variables with fewer measurement errors and, ultimately, more consistent
36
37 estimations. Moreover, even though outside collaborators are included here as a
38
39 regressor, our method of measuring regional networks imposes implicitly artificial
40
41 boundaries on the administrative limits of the region, thereby failing to identify possible
42
43 strong links between inventors located in separate regions, or links between co-located
44
45 inventors via a third inventor located outside the region. Suitable variables reflecting the
46
47 structure of the whole network and able to capture these two phenomena might alter
48
49 some of the results. However, additional theoretical developments should be carried out
50
51 before proceeding along these lines.
52
53
54
55
56
57
58
59
60

1
2
3 Next, as regards the KPF framework, a theoretical model of spatial interactions occurring
4 across regions in the production of knowledge, on which to base the estimated model is
5 lacking in this paper and, arguably, throughout the literature of spatial econometrics and
6 knowledge spillovers. Analogously, a specific, more meaningful modelling of these
7 interactions in the spatial weights matrix is needed. Fortunately, co-patenting and mobility
8 data are opening up a promising line of research in this direction, on which we will no
9 doubt focus our attention sooner rather than later. Moreover, lagging variables of the right
10 hand side of the models seeks to reduce the problems of endogeneity and reverse
11 causality, thus implicitly assuming that weak exogeneity is sufficient to obtain consistent
12 results. However, we should acknowledge the possibility that patenting activity in earlier
13 years may well have influenced the levels of certain independent variables, and therefore
14 consistency will be affected. To the best of our knowledge, suitable instruments have still
15 to be found for the explanatory variables, and so further research along these lines must
16 first be undertaken.¹⁷
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 Acknowledgements

41
42 Part of this work was carried out while Ernest Miguélez was visiting the Kiel Institute for
43 the World Economy (Kiel, Germany). The use of the Institute's facilities is gratefully
44 acknowledged. The authors would also like to share their appreciation of helpful
45 comments received from Wan-Hsin Liu, Aïda Solé-Auró, and the participants of the
46 "XXXIV Simposio de la Asociación Española de Economía" (Valencia, December 10th-
47 12th 2009). The work made by Ismael Gómez-Miguélez is sincerely appreciated. We also
48 acknowledge financial support from the Ministerio de Ciencia e Innovación, ECO2008-
49
50
51
52
53
54
55
56
57
58
59
60

05314 and Ernest Miguélez, from the Ministerio de Educación, AP2007-00792.

However, any mistake or omission remains ours.

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;

1
2
3 ANSELIN L., VARGA A. and ACS Z. (1997) Local Geographic Spillovers between
4
5 University Research and High Technology Innovations, *Journal of Urban Economics* 42,
6
7 422-48;
8

9
10
11 ANSELIN L. and KELEJIAN H. H. (1997) Testing for spatial error autocorrelation in the
12
13 presence of endogenous regressors, *International Regional Science Review* 20, 153-82;
14
15

16
17
18 AUDRETSCH D. B. and FELDMAN, M. (2004) Knowledge Spillovers and the Geography of
19
20 Innovation, in V. HENDERSON and J. THISSE (Eds) *Handbook of Urban and regional*
21
22 *Economics*, Volume 4;
23
24

25
26
27 AUTANT-BERNARD C. and LESAGE J. (2011) Quantifying Knowledge Spillovers using
28
29 Spatial Econometric Models, *Journal of Regional Science* 51, 471-96;
30
31

32
33
34 AUTANT-BERNARD C., MAIRESSE J. and MASSARD N. (2009) Spatial knowledge diffusion
35
36 through collaborative networks. Introduction, *Papers in Regional Science* 86, 341-50;
37
38

39
40 BALCONI M., BRESCHI S. and LISSONI F. (2004) Networks of inventors and the role of
41
42 academia: an exploration of Italian patent data, *Research Policy* 33, 127-45;
43
44

45
46
47 BERGMAN E. M. (2009) Embedding network analysis in spatial studies of innovation, *The*
48
49 *Annals of Regional Science* 43, 559-65;
50
51

52
53
54 BETTENCOURT L., LOBO J., HELBING D., KÜHNERT C., and WEST G. (2007a) Growth,
55
56 innovation, scaling, and the pace of life in cities, *Proceeding of the National Academy of*
57
58 *Sciences* 104, 7301-06;
59
60

1
2
3
4
5 BETTENCOURT L., LOBO J., and STRUMSKY D. (2007b) Invention in the city: Increasing
6
7 returns to patenting as a scaling function of metropolitan, *Research Policy* 36, 107-20;
8
9

10
11 BODE E. (2004) The spatial pattern of localized R&D spillovers: an empirical
12
13 investigation for Germany, *Journal of Economic Geography* 4, 43-64;
14
15

16
17
18 BOSCHMA R. (2005) Proximity and Innovation: A Critical Assessment, *Regional Studies*
19
20 39, 61-74;
21
22

23
24
25 BOSCHMA R., ERIKSSON R. and LINDGREN U. (2009) How does labour mobility affect the
26
27 performance of plants? The importance of relatedness and geographical proximity,
28
29 *Journal of Economic Geography* 9, 169-90;
30
31

32
33
34 BOTTAZZI L. and PERI G. (2003) Innovation and spillovers in regions: Evidence from
35
36 European patent data, *European Economic Review* 47, 687 - 710;
37
38

39
40
41 BRESCHI S. and CATALINI C. (2009) Tracing the linkages between science and
42
43 technology: An exploratory analysis of the research networks among scientists and
44
45 inventors, *Research Policy* 39, 14-26;
46
47

48
49
50 BRESCHI S. and LISSONI F. (2004) Knowledge networks from patent data: methodological
51
52 issues and research targets, in MOED H., GLÄNZEL W. and SCHMOCH U. (Eds)
53
54 *Handbook of Quantitative Science and Technology Research: The Use of Publication*
55
56 *and Patent Statistics in Studies of S&T Systems*, pp. 613-643. Springer Verlag, Berlin;
57
58
59
60

1
2
3 BRESCHI S. and LISSONI F. (2006) Cross-firm inventors and social networks: localised
4
5 knowledge spillovers revisited, *Annales d'Economie et de Statistique*, 79-80;
6
7

8
9 BRESCHI S. and LISSONI F. (2009) Mobility of skilled workers and co-invention networks:
10
11 an anatomy of localized knowledge flows, *Journal of Economic Geography* 9, 439-68;
12
13

14
15 CORREDOIRA R. A. and ROSENKOPF L. (2006) Learning from those who left: the reverse
16
17 transfer of knowledge through mobility ties, *Management Department Working Paper*;
18
19

20
21 COWAN R. and JONARD N. (2004) Network structure and the diffusion of knowledge,
22
23 *Journal of Economic Dynamics and Control* 28, 1557-75;
24
25

26
27
28
29 CRESPI G., GEUNA A. and NESTA L. (2007) The mobility of university inventors in
30
31 Europe, *Journal of Technology Transfer* 32, 195-215;
32
33

34
35
36 DE CLERCQ D. and DAKHLI M. (2004) Human capital, social capital, and innovation: A
37
38 multi-country study, *Entrepreneurship and Regional Development* 16, 107-28;
39
40

41
42 DE DOMINICI L., FLORAX R. J. G. M. and DE GROOT H. L. F. (2007) Regional clusters of
43
44 economic activity in Europe: Are social and geographical proximity the key
45
46 determinants?, *Knowledge for Growth - Industrial Research & Innovation (IRI)*;
47
48

49
50
51 DÖRING, T. and SCHNELLENBACH, J. (2006) What do we know about geographical
52
53 knowledge spillovers and regional growth?: A survey of the literature, *Regional Studies*
54
55 40, 375-95;
56
57
58
59
60

1
2
3 FINGLETON B. and LE GALLO J. (2008) Estimating spatial models with endogenous
4
5 variables, a spatial lag and spatially dependent disturbances: Finite sample properties,
6
7 Papers in Regional Science 87, 319-39;
8
9

10
11 FLEMING L., KING C., and JUDA A. (2007) Small worlds and innovation, Organization
12
13 Science 14, 375-93;
14
15

16
17
18 GOMES-CASSERES B., HAGEDOORN J., and JAFFE A. B. (2006) Do alliances promote
19
20 knowledge flows?, Journal of Financial Economics 80, 5-33;
21
22

23
24
25 GRANOVETTER M. S. (1985) The strength of weak ties, American Journal of Sociology
26
27 78, 1360-80;
28
29

30
31
32 GRILICHES Z. (1979) Issues in assessing the contribution of research and development to
33
34 productivity growth, The Bell Journal of Economics 10, 92-116;
35
36

37
38 GRILICHES Z. (1991) Patent Statistics as Economic Indicators: A Survey, NBER Working
39
40 Papers No. 3301, National Bureau of Economic Research;
41
42

43
44
45 GROSSMAN. G. M. and E. HELPMAN (1991) Innovation and Growth in the Global
46
47 Economy. MIT Press, Cambridge, MA;
48
49

50
51
52 HARRIS R. and KRAVTSOVA V. (2009) In Search of 'W', SERC Discussion Paper 17,
53
54 London;
55
56

1
2
3 HAUSMAN J.A., HALL B.H. and GRILICHES Z. (1984) Econometric models for count data
4
5 with an application to the patent-R&D relationship, *Econometrica* 51, 909-38;
6
7

8
9
10 HOEKMAN J., FRENKEN K. and VAN OORT F. (2009) The geography of collaborative
11
12 knowledge production in Europe, *The Annals of Regional Science* 43, 721-38;
13

14
15
16 HOISL K. (2007) Does mobility increase the productivity of inventors?, *Journal of*
17
18 *Technology Transfer* 34, 212-25;
19

20
21
22 HOISL K. (2009) Tracing mobile inventors: The causality between inventor mobility and
23
24 inventor productivity, *Research Policy* 36, 615-36;
25
26

27
28
29 JAFFE A. B. (1986) Technological opportunity and spillovers of R&D: evidence from
30
31 firms patents, profits and market value, *American Economic Review* 76, 984-1001;
32
33

34
35
36 JAFFE A. B. (1989) Real effects of academic research, *American Economic Review* 79,
37
38 957-70;
39

40
41
42 KELEJIAN, H. H. and PRUCHA, I. R. (1998) A generalized spatial two-stage least squares
43
44 procedure for estimating a spatial autoregressive model with autoregressive disturbances,
45
46 *Journal of Real Estate Finance and Economics* 17, 99-121;
47
48

49
50
51 KELEJIAN H. H. and PRUCHA I. R. (2007) HAC estimation in a spatial framework,
52
53 *Journal of Econometrics* 140, 131-54;
54
55

1
2
3 KELEJIAN H. H., PRUCHA I. and YUZEFOVICH Y. (2004) Instrumental Variable
4
5 Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large
6
7 and Small Sample Results, in J. LESAGE and R. K. PACE (Eds) Spatial and Spatiotemporal
8
9 Econometrics, Advances in Econometrics, Vol. 18, pp. 163-198. Elsevier, New York;

10
11
12
13
14 KENNEDY, P. (1992) A guide to econometrics, 3rd Edition. Blackwell, Oxford;

15
16
17
18 KIM J., LEE S. J. and MARSCHKE G. (2006) International Knowledge Flows: Evidence
19
20 from an Inventor-Firm Matched Data Set, NBER Working Paper No. 12692;

21
22
23
24 KROLL H. (2009) Spillovers and proximity in perspective: A network approach to
25
26 improving the operationalisation of proximity, Working Papers Firms and Regions No.
27
28 R2/2009, Fraunhofer ISI;

29
30
31
32
33 Laudel G. (2003) Studying the brain drain. Can bibliometric methods help?,
34
35 *Scientometrics* 57, 215-37;

36
37
38
39
40 LENZI C. (2009) Patterns and determinants of skilled workers' mobility: evidence from a
41
42 survey of Italian inventors, *Economics of Innovation and New Technology* 18, 161-79;

43
44
45
46
47 LOBO J. and STRUMSKY D. (2008) Metropolitan patenting, inventor agglomeration and
48
49 social networks: A tale of two effects, *Journal of Urban Economics* 63, 871-84;

50
51
52
53
54 LUCAS R. E. (1988) On the mechanics of economic development, *Journal of Monetary*
55
56 *Economics* 22, 3-42;

1
2
3 MAGGIONI M. A. and UBERTI T. E. (2008) Knowledge networks across Europe: which
4 distance matters?, *The Annals of Regional Science* 43, 691-720;

5
6
7
8
9 MIGUÉLEZ E. and GÓMEZ-MIGUÉLEZ I. G. (2010) Singling out individual inventors from
10 patent data, *IREA Working Paper* 2011/05;

11
12
13
14
15
16 MIGUELEZ E., MORENO R. and ARTÍS M. (2011) Does social capital reinforce
17 technological inputs in the creation of knowledge? Evidence from the Spanish regions,
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Regional Studies 45, 1019-38;

25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
MORENO R., PACI R. and USAI S. (2005a) Spatial spillovers and innovation activity in
European regions, *Environment and Planning A* 37, 1793-812;

31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
MORENO R., PACI R. and USAI S. (2005b) Geographical and sectoral clusters of
innovation Europe, *Annals of Regional Science* 39, 715-39;

38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
PERI G. (2005) Determinants of Knowledge Flows and Their Effect on Innovation, *The
Review of Economics and Statistics* 8, 308-22;

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
PONDS R., VAN OORT F. and FRENKEN K. (2007) The geographical and institutional
proximity of research collaboration, *Papers in Regional Science* 86, 423-43;

51
52
53
54
55
56
57
58
59
60
PONDS R., VAN OORT F. and FRENKEN K. (2010) Innovation, spillovers and university-
industry collaboration: an extended knowledge production function approach, *Journal of
Economic Geography* 10, 231-55;

1
2
3 RODRÍGUEZ-POSE A. and CRESCENZI R. (2008) Research and Development, Spillovers,
4
5 Innovation Systems, and the Genesis of Regional Growth in Europe, *Regional Studies* 42,
6
7 51-67;
8

9
10
11 ROMER P. M. (1986) Increasing Returns and Long-run Growth, *Journal of Political*
12
13 *Economy* 94, 1002-37;
14

15
16
17 ROMER P. M. (1990) Endogenous Technological Change; *Journal of Political Economy*
18
19 98, 71-102;
20
21

22
23
24 ROSENKOPF L. and ALMEIDA P. (2003) Overcoming Local Search Through Alliances
25
26 and Mobility, *Management Science* 49, 751-66;
27
28

29
30
31 SAXENIAN A. (1994) *Regional Advantage. Culture and Competition in Silicon Valley and*
32
33 *Route 128*. Harvard University Press, Cambridge, MA;
34
35

36
37
38 SINGH J. (2005) Collaborative Networks as Determinants of Knowledge Diffusion
39
40 Patterns, *Management Science* 51, 756-70;
41
42

43
44 SINGH J. AND AGRAWAL A. K. (2011) Recruiting for Ideas: How Firms Exploit the Prior
45
46 Inventions of New Hires, *Management Science* 57, 129-50;
47
48

49
50 SHALEM R. and TRAJTENBERG M. (2008) Software patents, inventors and mobility,
51
52 Working Paper;
53
54

55
56 SONG J., ALMEIDA P. and WU G. (2003) Learning-by-hiring: When is mobility more to
57
58 facilitate interfirm knowledge transfer?, *Management Science* 49, 351-65;
59
60

1
2
3
4
5 TER WAL A. L. J. and BOSCHMA R. (2009) Applying social network analysis in economic
6 geography: framing some key analytic issues, *The Annals of Regional Science* 43, 739-56;
7
8

9
10
11 TRIPPL M. and MAIER G. (2010) Knowledge spillover agents and regional development,
12 *Papers in Regional Science* 89, 229-33;
13
14

15
16
17 THOMA G. and TORRISI S. (2007) Creating Powerful Indicators for Innovation Studies
18 with Approximate Matching Algorithms. A test based on PATSTAT and Amadeus
19 databases, *CESPRI Working Papers*, No. 211;
20
21
22
23

24
25
26 THOMA G., TORRISI S., GAMBARDILLA A., GUELLEC D., HALL B.H., and HARHOFF D.
27 (2009a) Methods and software for the harmonization and combination of datasets: A test
28 based on IP-related data and accounting databases with a large panel of companies at the
29 worldwide level, mimeo;
30
31
32
33
34

35
36
37 THOMA G., TORRISI S., GAMBARDILLA A., GUELLEC D., HALL B.H., and HARHOFF D.
38 (2009b) Harmonizing and Combining Large Datasets: An Application to Firm-level
39 Patent and Accounting Data, *STI Working Paper*, OECD, Paris;
40
41
42
43
44

45
46
47 WASSERMAN S. and FAUST K. (1994) *Social Network Analysis: Methods and*
48 *Applications*. Cambridge University Press, Cambridge.
49
50
51
52
53
54
55
56
57
58
59
60

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(CONN_ABS) | 0.08*** (0.02) | | | 0.12*** (0.02) | | |
| ln(CONN_INV) | | | 0.16*** (0.04) | | | 0.22*** (0.05) |
| ln(SIZE_LC) | | 0.03 (0.04) | | | 0.08* (0.04) | |
| 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(CONN_ABS) | 0.07*** (0.02) | | | 0.11*** (0.02) | | |
| ln(CONN_INV) | | | 0.13*** (0.04) | | | 0.21*** (0.04) |
| ln(SIZE_LC) | | 0.06* (0.04) | | | 0.08** (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(CONN_ABS) | 0.07*** (0.02) | | | 0.11*** (0.02) | | |
| ln(CONN_INV) | | | 0.13*** (0.04) | | | 0.21*** (0.04) |
| ln(SIZE_LC) | | 0.06* (0.04) | | | 0.08** (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(CONN_ABS) | 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).

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 |
| Absolute connectivity | Absolute number of edges | 1993-2002 | REGPAT |
| Relative connectivity | Number of edges to no. of inventors | 1993-2002 | REGPAT |
| Size of largest component | % inventors in the largest component | 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 ²) | 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 |

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

¹ We would like to acknowledge the referees for raising this point.

² 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.

³ 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.

⁴ The number of these 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.

⁵ The R&D productivity index in the knowledge production function would be equivalent to the Total Factor Productivity index in a production function.

⁶ 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).

⁷ We use the first and second order spatial lags, as suggested in KELEJIAN et al. (2004).

⁸ 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.

⁹ 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.

¹⁰ International Patent Classification

¹¹ The technological specialisation index is designed to control for the existence of Marshall-Arrow-Romer (MAR) externalities derived from the regional specialisation 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.

¹² Although overall employment in these sectors would be a better proxy, these data are not available.

¹³ 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.

¹⁴ 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.

¹⁵ To see these points, the quadratic form of population density was also included in all the models. By doing so, the population density variable becomes positive and significant, whilst its quadratic form remains negative. These results would therefore confirm the existence of congestion effects. For the sake of brevity, the results are not reported here but can be provided upon request from the authors.

¹⁶ 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.

¹⁷ 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

1
2
3
4 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
5
6 regressor. The results - available upon request - make few changes to our conclusions, with the exception
7
8 of the “outside collaborators” variable, which was no longer significant at 10%.
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review Only