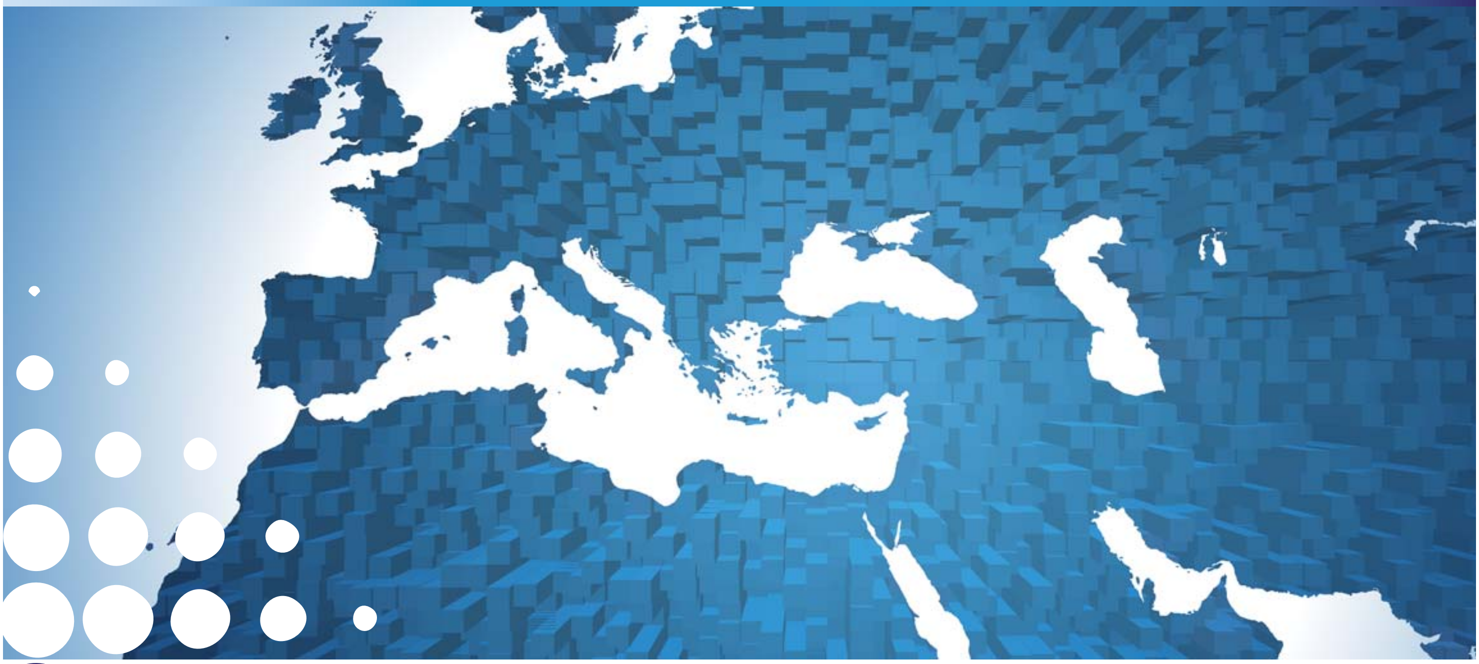
WP4/01 SEARCH WORKING PAPER

Knowledge production function and proximities. Evidence from spatial regression models for the European regions.

Emanuela Marrocu, Raffaele Paci, Stefano Usai

January 2013



SEARCH
SHARING KNOWLEDGE ASSETS: INTERREGIONALLY COHESIVE NEIGHBORHOODS

EUROPEAN COMMISSION
European Research Area

SEVENTH FRAMEWORK PROGRAMME

Funded under Socio-economic Sciences & Humanities

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2010-2.2-1) under grant agreement n° 266834

Knowledge production function and proximities.
Evidence from spatial regression models for the European regions.

Emanuela Marrocu, Raffaele Paci, Stefano Usai

University of Cagliari, CRENoS, Italy

Abstract

This paper aims at investigating the connections among regional innovation systems along several proximity dimensions. In particular, we assess if, and how much, the creation of new ideas in a certain region is the result of internal efforts as much as of knowledge flows coming from other regions which may be considered neighbors not only in the geographical space but also in the institutional, technological, social and organizational one. The analysis, based on spatial econometric techniques, is implemented for an ample dataset referring to 276 regions in 29 European countries (EU27 plus Norway, Switzerland) for the last decade.

Results attest the importance of exploiting relationships with spatially close-by regions but also with technologically proximate regions which may prove geographically very distant. This implies that, no matter whether the knowledge flow is due to collaborations, imitation or workers mobility, there is a global perspective which goes beyond the regional and the national systems of innovation. Global networks connecting regions which share a similar specialization pattern and therefore the same cognitive and technological base may be as crucial as local cluster of regions in enhancing technological advance. On the whole, regional system of innovations have to be effective not only in exploiting their internal production of ideas based on R&D investments and human capital but also in creating synergies with other regions to absorb part of their knowledge and expertise.

This interpretative scenario is crucial for regions in the European Neighboring Countries since it enlarges the potential basin of knowledge externalities which may help their technological catching up with respect to advanced regions in the European Union.

Keywords: technological production, proximity, networks, human capital
JEL: O31, R12, O18, C31

Acknowledgments: We would like to thank Barbara Dettori and Marta Foddi for excellent research assistance. This report is based on materials developed in Marrocu, Paci and Usai (2011 and 2012).

1. Introduction

The current economic downturn is forcing countries and regions to design policies which are able to balance short and long run effects in the most effective way, while saving resources. The European Union is trying to achieve such goals with a complex set of interventions where the European innovation strategy, as set out in the Innovation Union document, is the crucial instrument to achieve sustainable and inclusive growth in the long run. This strategy, essentially, targets the ability of each region to improve its internal and, most importantly, external links, since regions need to confront themselves with the worldwide scenario at large to improve their connections and cooperation with other territories, clusters and innovation players.

It is widely recognised that the capacity of a region to generate, transmit and acquire knowledge and innovation depends on a multifaceted set of factors: investment in R&D, work force experience, education and training, collaboration networks, technology transfer mechanisms, researchers' and workers' mobility. In particular, the literature has distinguished between the creation of new ideas and inventions and the absorption of innovations generated in other regions. Several works, both on theoretical (Grossman and Helpman, 1990; Rallet and Torre, 1999, Antonelli, 2008) and empirical grounds (Jaffe, 1989; Coe and Helpman, 1995), have argued that innovation depends on investments in research and human capital as much as on interactive learning, knowledge diffusion and circulation of ideas.

Both sets of aspects are strictly related to the concept of closeness of economic agents and how proximity affects their ability to connect and, possibly, cooperate within systemic networks at different territorial levels. The concept of closeness has several dimensions and may have different implications; obviously, the most common one refers to geography: spatial concentration is believed to be crucial in the dynamics of innovation, thanks mainly to local spillovers. However, local relations go often together with wider links and networks. In this respect, the spatial dimension may be just a counterpart of other forms of a-spatial proximity: institutional, cognitive or technological, social or relational and organizational, as exhaustively argued and commented by Boschma (2005).

In this perspective, the general object of this paper is to analyse how internal and external factors interact in determining the technological performance of the European territories. More specifically, we investigate to what extent the regional inventive activity depends on intra-regional characteristics (mainly R&D expenditure and human capital) and on regions' ability to absorb inter-regional knowledge spillovers channelled and diffused by different types of proximity. These aspects are investigated by applying spatial econometric techniques to a Knowledge Production Function (KPF) model. With respect to its traditional formulation, this is augmented with extraregional factors, mediated by different kinds of proximity and networks (institutional, geographical, technological, social and organizational), which are expected to enhance a region's innovative activity. Our analysis is based on an ample dataset referring to 276 regions in 29 countries (EU27 plus Norway and Switzerland) over the last decade. The particular richness of the information required for the implementation of our econometric analysis denies the possibility to extend our sample to regions in the European Neighbouring Countries. Nevertheless, results are sufficiently broad to allow for a generalisation of some far reaching conclusions.

The regional scenario of the enlarged Europe examined in the paper represents an extremely interesting case study, as the high heterogeneity in terms of economic as well as innovative regional performance (Hollander et al., 2009) asks for coordinated policy interventions both at the national and regional level. Such interventions, defined in the Research and Innovation Strategies for Smart Specialisation (RIS3 strategies) document, are intended to provide a coherent national and regional framework to ensure knowledge development. The paper, therefore, addresses the following main research questions:

1) what is the balance of internal and external factors in shaping regional innovative performance?

2) what kind of connections are most effective in driving knowledge spillovers across regions?

3) Are these connections complementary or substitute?

4) What lessons can be gained from such results to design more effective innovation policies both for EU and ENC regions?

The main original contribution of this paper is, therefore, to assess empirically the joint and complementary effectiveness of five different dimensions of proximity – geographical, institutional, technological, social and organizational – in channelling knowledge spillovers at the aggregate regional level in Europe.

Our model selection strategy points out that a spatial autoregressive specification is adequate to model the regional interconnectivity pattern which enables the transmission of knowledge across proximate territories. In order to fully account for the complementarities among all the proximity dimensions considered, the optimal estimation strategy would entail the specification of a comprehensive model which includes all of them at the same time. However, for a spatial autoregressive specification, this goes beyond the current state of the art (Elhorst, 2010) as it would require the formulation of new econometric tools. As a workable alternative we, therefore, adopt the SAR model variant which includes two spatially lagged terms for the dependent variable, computed on the basis of two different proximity dimensions; the institutional proximity is considered throughout the models by including a full set of country dummies. This kind of specification, which allows to account for one pair of proximity dimensions at a time, was first proposed in a different setting by Lacombe (2004). Moreover, to obtain an approximate measure of the overall multiplier effect resulting from the simultaneous working of all knowledge transmission mechanisms, we carry out a post-estimation exercise based on model combining techniques.

Our main results confirm the importance of investment in R&D and reveal that human capital plays an even greater role in fostering innovative activity and in generating inflows of knowledge relevant to region's existing knowledge base. More importantly, our analysis shows that geography is not the only dimension which may favour knowledge diffusion and not even the most important one. Technological proximity proves to be the most relevant, while social and organizational networks are also significant although their role is modest. This implies that policy interventions have to be coordinated with several different instruments and along diverse dimensions to be effective in reaching the overall innovation targets.

The paper is structured as follows. In the next section we present concepts, definitions and measures of the different channels of technological externalities which correspond to the five proximity dimensions. In the third section we describe the empirical model of the Knowledge Production Function which is later implemented with a specific estimation strategy in the fourth section. The fifth section discusses the empirical results of the several models which have estimated. In section 6 we draw some concluding observations while the relevant policy implications are discussed in section 7.

2. Proximity dimensions: concepts and measures

The idea that technological progress is a complex process which combines the direct production of innovation at the local level together with the absorption of the knowledge produced in the global setting is by now widely shared. Economic literature from different schools of thoughts provides theoretical backing to this idea, which is based on the presence of local spillovers both within and across regions and countries (see Castellacci, 2007 and Christ, 2009, for recent surveys). Such spillovers are obviously related to the geographical dimension since close-by agents are believed to have a better innovative performance because of pecuniary and pure technological advantages[1]. More specifically, they have less costly access to information and they can share tacit knowledge (a local public good) through face to face contacts. Nonetheless, the French School of

[1] Antonelli et al. (2008) argue that there may be an optimal size of local knowledge pools since a low innovation network density reduces access to external knowledge whilst an excessively large one enhances congestion and reduces appropriability.

Proximity (Kirat and Lung, 1999; Torre and Gilly, 2000) argues that geographical proximity is neither necessary nor sufficient and that there may be a separate role for a-spatial links among economic entities (see Carrincazeaux and Coris, 2011, for a recent review). The exchange of knowledge and technological interdependence, in other words, may be related, according to Boschma (2005), to at least four other dimensions of proximity across agents: institutional, technological (or cognitive), social (or relational) and organizational.

2.1 Definitions and previous literature

In this section we provide a definition of these four concepts of proximity and a description of their measurement as suggested in the empirical literature.

Institutional proximity means that the effective transmission of knowledge may be facilitated by the presence of a common institutional framework. Institutions, such as laws and norms, can provide a set of standard procedures and mechanisms which are shared by agents and, therefore, taken for granted. This mutual endowment proves relevant in reducing uncertainty and in lowering transaction costs and, thus, favours cooperative behaviours in the regional context (Maskell and Malmberg, 1999; Gertler, 2003).

Technological (or cognitive) proximity indicates that knowledge transfer requires specific and appropriate absorptive capacity (Cohen and Levinthal, 1990), which entails, among others, a homogenous cognitive base with respect to the original knowledge in order to understand and process the new incoming knowledge effectively[2]. In practical terms, we expect that economic agents who share a similar knowledge base, or territories which have in common a similar specialisation structure, can exchange information more easily and less costly, and this may favour innovation.

Social (or relational) proximity refers to the fact that economic relationships may reflect social ties and vice versa (Granovetter, 1985). In the context of innovation processes, this implies that social closeness facilitates firms' capacity to learn, absorb external knowledge and innovate since social nearness breeds trust which, in turn, lowers transaction costs and facilitates collaboration. This aspect can be particularly relevant for a risky and uncertain phenomenon such as technological progress.

Organisational proximity refers to the relations within the same group or organisation which influence the individual capacity to acquire new knowledge coming from different agents. It

[2] The concept of absorptive capacity does not depend only on cognitive proximity and has a wider application at the level of firms, sectors, regions and nations. In particular, Iammarino (2005) observes that the ability of a region to absorb and generate new knowledge depends on skills which are people- and institution-embodied, that is human capital and R&D investments.

reduces uncertainty and incentives to opportunistic behaviour since it provides an area of definition of practices and strategies within a set of rules based on organizational arrangements (Kirat and Lung, 1999). Such arrangements can be either within or among firms and may take different forms along a range which goes from informal relations among companies to formally organised firms.

The different dimensions of proximity discussed above can be seen as a crucial condition for firms' interaction and cooperation aimed at innovation. Boschma and Frenken (2010), in particular, explain how proximity (or similarity) can act as a driving force for the formation and the evolution of networks. The interconnected role of proximity and networks on local innovation performance can be analysed thanks to the KPF approach, introduced by Griliches (1979) to study the relationship between knowledge inputs and outputs at the firm level. Since then it has been extensively used to analyse how this relationship works both at the firm and at the territorial level. In particular, regional KPFs have been estimated to assess the role of both internal and external factors on regional innovation systems. The seminal paper by Jaffe (1989), who proves the existence of geographically mediated spillovers from university research to commercial innovation in US metropolitan areas. The main results of his paper have been later extended and strengthened by many other authors who provide evidence in favour of local externalities both within and across regions in the USA (Acs et al., 1992; Anselin et al., 1997; O´hUallacha´in and Leslie, 2007). Most of these studies introduce the concept of geographical proximity and test its importance by means of spatial econometric techniques.

Along the same vein, several studies have been proposed for the EU regions (Tappeiner et al., 2008; Acosta et al., 2009; Buesa et al., 2010 are among the latest contributions).[3] These studies find that innovation performance is partly due to internal factors and partly to spillovers which flow from one region to another. Unlike the USA studies, some of these articles add other possible dimensions of proximity and assess their role on knowledge production. In particular, Bottazzi and Peri (2003), Greunz (2003) and Moreno et al. (2005) investigate inter-regional knowledge spillovers across European regions, testing whether technological proximity influences the creation of new knowledge within European regions. Results show that interregional knowledge spillovers exist both between close-by regions and between distant regions with similar technological profiles. This indicates that geographical distance is not the only dimension to be investigated and that knowledge spillovers may be induced also by cognitive closeness. Furthermore, all these studies consider institutional proximity (measured by means of country dummies) and find it relevant in indentifying the more and less innovative regions.

[3] The only contributions which analyze different continents at the regional level are Crescenzi et al. (2007) for US and EU, with data coming from USPTO and EPO respectively, and Usai (2011) on OECD regions with homogenous information on Patent Cooperation Treaty applications.

Only few contributions examine the role of social or relational networks[4] together with geographical proximity within a KPF[5]. Maggioni et al. (2007), Kroll (2009) and Ponds et al. (2010) find that both the local neighbourhood and the connections with other regions based on co-operation matter for the local process of knowledge generation. The first article measures social proximity by means of cooperation networks for the Fifth Framework Programme, the second one uses co-patenting across regions, whilst the third uses co-publications. Other contributions have introduced various features of inventors' network in a KPF framework: Lobo and Strumsky (2008) for the case of the USA, MSA's and Miguelez and Moreno (2011) for the European NUTS2 regions. They all find that the scale and extent of networks have a positive impact on innovative performance. However, none of these studies operationalizes this concept in order to gauge proximity for each couple of regions[6], but rather they use it as a regional indicator which measures the region's degree of connectivity and openness.

Finally, to the best of our knowledge, there are no contributions which focus directly on the role of organizational proximity on regional innovation performance. The only partial exception is the article by Sorensen et al. (2006), where organizational proximity is considered as a determinant of knowledge flows proxied by citations. The use of micro data allows introducing organizational proximity as a binary variable which is equal to unity when the citation comes from employees of the same firm, even when they reside in different regions. Another interesting study on the impact of organizational proximity on innovation, although at the firm level, is Oerlemans and Meeus (2005), who, using survey-based micro data on the Netherlands, conclude that interregional relations with business agents (users and suppliers) lead to a better innovative performance.

2.2 Proximity measures at the regional level

In this section we analyse in detail how we operationalize the five concepts of proximities presented above into measures to be used later in the KPF estimation. For each dimension we try to clarify the mechanisms which link the micro level (agents, firms) where the closeness measures operate and the aggregate regional level which is investigated in the paper.

[4] Social proximity has been also included in studies of R&D cooperation networks, such as that of Autant-Bernard et al. (2007), who find that the probability of collaboration is influenced by each individual's position within the network and also that social distance seems to matter more than geographical distance. In the same vein, Hoekman et al. (2009) find negative effects of both geographical and institutional distance on research collaboration, using data on inter-regional research collaboration measured by scientific publications and patents in Europe.

[5] An interesting parallel study which has tried to provide a measure of different proximities, namely relational, social and technological, to assess their role in affecting productivity growth, rather than innovative activity, has been recently proposed by Basile et al. (2012).

[6] It is worth noting that Rodriguez-Pose and Crescenzi (2008) use the concept of 'social filter', a composite index describing the socio-economic realm of each region, in their study of regional growth in Europe. Moreover, the role of the social filter is assessed not only within regional borders but also across regions. This external role is, however, mediated only along the geographical dimension.

All proximity measures considered in this section are computed at the NUTS (Nomenclature des Unités Territoriales Statistiques) 2 level (see Appendix 1).[7] Although we are aware that some proximity notions, as it is particularly the case for the social and organizational ones, were initially formulated for firm level kind of connections, we think that if a significant impact is found, even at an aggregate territorial scale, this is to be interpreted as evidence in favour of the existence of underlying micro mechanisms, which are effective and pervasive in driving knowledge creation across regions. As for the social and organizational proximity, although they are both based on the network notion it is worth remarking that accounting for the sources of network creation and development goes beyond the scope of this paper. Following the traditional regional science literature, we consider both social and organizational structures to be the result of fixed or slowly evolving networks (Corrado and Fingleton, 2012).

In Table 1 we summarize the main descriptive statistics of the proximity measures considered, along with their correlation matrix while the data sources and definitions are reported in Appendix 2.

Geographical proximity. This is the standard and widely used indicator of proximity, it is measured by the distance in km between the centroids of any two regions. This measure is preferred to the contiguity matrix since it allows one to consider all the potential interactions among regions so that spillovers are not limited to those regions which share a border. The median spatial distance across regions in Europe is 1270 km, ranging from a lowest value of 18 km among Belgium's regions to the maximum distance, that is 4574 km, between Cyprus and Ireland. In the econometric analysis we use the inverse of the distance so that high values indicate more proximate regions and thus a higher probability to exchange knowledge. Moreover, we assess which is the most relevant distance range in determining knowledge spillovers.

Institutional proximity. Knowledge is transmitted more easily when individuals and firms share the same institutional framework, a common language and similar cultural, ethnic and religious values. Thus two regions belonging to the same national institution are expected to have higher knowledge exchange. A simple, and widely used, way to account for these time invariant common factors is to include a full set of country dummies. Alternatively, we model institutional proximity by means of a weight matrix, whose elements take value 1 if two regions belong to the

[7] For the small European countries (Cyprus, Estonia, Lithuania, Luxembourg, Latvia, Malta) the regional breakdown is not available so they are considered at the country (NUTS0) level. Although we acknowledge that the NUTS2 territorial scale may be too aggregate to unveil all potential spillovers, nonetheless it is the observational level for which consistent regional data is made available by statistical offices, thus enabling us to consider the widest possible coverage of the European territory.

same country and zero otherwise.[8] We anticipate here that the empirical specification based on such a proximity matrix is outperformed by the estimation which includes country dummies to account for the importance of institutional similarity across regions. Note also that the inclusion of national indicators is also suitable to account for the potential adverse influence due to "border effects"; the international trade literature (Anderson and Wincoop, 2004) has largely emphasized how such effects may inhibit trade among countries, and this can analogously happen in this case of knowledge flows when the regional and the national boundaries coincide (Parent and LeSage, 2012).

Technological proximity. In order to attract new knowledge from outside, firms and regions may need to build up absorptive capacity around the existing knowledge base and carry out technological activity in similar fields. In other words, cognitive capacity is bounded and companies and regions sharing an analogous knowledge base may exchange information and knowledge and learn from each other more easily. To measure the technological, or cognitive, proximity across regions we compute a similarity index between region *i* and region *j*, based on the distribution of patenting activity among 44 sectors,[9] defined as:

$$t_{ij} = 1 - \left(\frac{1}{2}\sum_{k=1}^{K=44}\left|l_{ik} - l_{jk}\right|\right)$$

where l_{ik} is the sectoral share of sector *k* in region *i*. The index t_{ij} is defined between zero (perfect dissimilarity of the sectoral distribution) and one (perfect similarity); thus, the higher the index value, the more similar the technological structure of the two regions and the higher the probability that they can exchange knowledge. The index is computed for each couple of regions to build up a technological proximity matrix *T* with generic element t_{ij}.[10] The 44 sectors are defined on the basis of patenting activity measured at 2-digit SIC level and they mainly refer to manufacturing industries where most of the patenting activity is performed.

In Table 1 we show that the two most technologically distant regions (Ionia Nisia and Notio Aigaio in Greece) exhibit an index of 0.05. Interestingly, the higher degree of technological

---

[8] A similar matrix is used by Paci and Usai (2009) to analyze how institutional factors positively affect the flows of knowledge for the case of EU15 regions.

[9] Compared to other studies our sectoral breakdown is quite fine and informative. For instance, Parent and LeSage (2012) consider 8 sectors in analyzing knowledge determinants for the whole economic regional system of nine Western European countries.

[10] We have also computed a matrix based on the correlation coefficient among the sectoral patent shares between regions *i* and *j* as in Jaffe (1986) and Moreno et al. (2005). The matrices based on the similarity and correlation coefficients are highly correlated (the sample correlation coefficient is 0.91) and they give very similar results; therefore in the following sections we present only the results based on the similarity index.

similarity (0.94) is found in two non-adjacent regions, located in different countries: Piedmont in Italy and Niederbayern in Germany. The econometric estimation allows to test whether regions with a similar technological specialization, for instance in high tech industries, and therefore with a common cognitive background are more likely to benefit from mutual knowledge flows, regardless of their geographical location.

In order to test the robustness of the technological proximity measure based on patenting activity we have also computed a matrix based on the sectoral distribution of employment, which is available for seventeen 2-digit NACE manufacturing and service sectors. However the matrix based on the finer distribution of patenting activity proves to be more able to grasp the informative content of the cognitive similarity among territorial units and therefore in the regression tables we report only the results based patenting activity.

Social proximity. The main idea is that individuals who have socially embedded relations and networks are more likely to trust each other and therefore to exchange tacit knowledge smoothly. At the macro level this implies that regions where network members reside are facilitated in exchanging knowledge. In this paper we measure social proximity by means of co-inventorship relations among multiple inventors of the same patent in case they are resident in different regions. As a result, the generic element s_{ij} of the symmetric social matrix S is defined as the number of inventors located in region i which have co-operated with inventors located in region j to conceive a patented invention. In this matrix we do not consider the intra-regional relationships, the principal diagonal elements are therefore set to zero. The rationale is that the number and the intensity of links among inventors located in different regions are able to catch the existence of a social network between regions which facilitates the exchange of knowledge.

Table 1 shows that the number of non-zero links (co-inventorships) in the matrix represents only a small fraction (18%) of all potential relationships. The highest social interaction (137) is reached by the two contiguous German regions of Düsseldorf and Köln, followed by other couples of contiguous German regions located in the industrialized area of Baden-Wurttemberg: Karlsruhe with Rheinhessen-Pfalz and Stuttgart with Karlsruhe. Thus, there is a geographically defined cluster of regions characterized by strong social relationships measured by co-inventorships. As expected, spatial proximity favours social interactions among inventors although, from Table 1, we can see that the correlation coefficient between the geographical and social proximity matrices is quite small (0.12).[11]

[11] It is interesting to notice that the correlation coefficient with the contiguity matrix is much higher (0.39), signaling that strong social relationships are more likely to develop among contiguous regions.

Organizational proximity. Organizational proximity refers to the connections within the same organization or group which explain the capacity of an agent to acquire knowledge coming from a multitude of different actors. For example, we can think of establishments belonging to the same firm, departments of the same university or employees working for the same company. We follow Picci (2010) and Maggioni et al. (2011) who measure proximity across nations and regions, respectively, by using the affiliation to the same organization of the applicant and the inventors of a patent. Given this definition, we are not considering the case in which the applicant and the inventor are the same as much as the case in which they are different but located in the same region. As a result the main diagonal is set to zero. A characteristic of the applicant-inventor matrix is that it is not symmetric. In other words, the relationships originated by the applicant in region i with inventors resident in region j are different with respect to the links between applicant in region j and inventors living in region i. Since we are interested in the total number of organizational relationships between the two regions, we sum up mirror cells so that the generic element o_{ij} of the organizational matrix O is defined as the total number of bilateral relationships between applicants and inventors located in the regions i and j.

As with the previous types of proximity, we expect a positive influence of organizational networks in the process of knowledge creation and diffusion since they are believed to reduce uncertainty and opportunism. Table 1 shows that the number of non-zero links in the organizational matrix amounts to 17% of total possible relationships among European regions. Interestingly, the highest value (480) is reached by two distant regions within France: Île de France and Rhône Alpes. The former hosts the capital, Paris, where most French companies locate their headquarters, whilst the latter is renowned for its scientific parks and research laboratories which are apparently linked to parent companies. In this case the hypothesis, to be tested empirically, is that the two regions are characterized by a high organizational proximity which should facilitate the knowledge exchange between them.

It is worth pointing out that both social and organization proximity measures are not completely satisfactory[12], the phenomena they are intended to capture are very complex and their measurement is a challenging task even at the micro level. However, we think that our contribution is, at least partly, successful attempt at responding to Anselin's (2010) solicitation for a more and more adequate representation of the spatial processes by deriving their interconnectivity structure on the basis of agents' social and economic interaction. This aspect is becoming increasingly relevant and deserves further investigation, in future analysis we intend to search for different

[12] As a matter of fact it is quite difficult to obtain a non-overlapping measure of organizational and social proximity. Indeed, the correlation coefficient between the two proximity matrices reported in Table 1 is 0.74.

proxies of social and organization closeness in order to reduce their overlapping and thus to gain a better understanding of their distinctive role in conducing knowledge flows.

3. The empirical KPF model

In this section we first present the econometric model used to investigate the determinants of the process of knowledge creation and diffusion in Europe, followed by a description of the data used for the dependent variable and for the production inputs considered.

The literature on the determinants of innovative activity at firms' and regional level has been traditionally based on the estimation of a KPF model, where the output is measured by the patenting activity and the input by the R&D expenditure. We follow this approach but we augment the KPF specification by introducing human capital as an additional input, given its well-known effects on knowledge creation. Indeed, in the case of traditional sectors and small enterprises, the creation of innovation is not necessarily the result of a formal investment in research but it is often derived either from an informal process of learning by doing (Nelson and Winter, 1982) or from the absorption of external knowledge (Abreu et al., 2008). Firms' and regions' ability to understand, interpret and exploit internal and external knowledge relies on prior experiences embodied in individual skills and, more generally, in a well-educated labour force (Engelbrecht, 2002 and Archibugi and Filippetti, 2011). In light of the discussion above, we also explicitly consider the presence of external factors coming from "proximate" regions, which may enhance the impact of the internal ones thanks to spillover effects.

Thus, the general form of the empirical model for the KPF is specified according to a log-linearized Cobb-Douglas production function as:

$$inn_i = \beta_1 rd_i + \beta_2 hk_i + \phi \ controls_i + \gamma \ \ proximity \ \ factors_i + \varepsilon_i \qquad (1)$$

where lower case letters indicate log-transformed variables. More specifically, the innovation output *inn* is proxied by the yearly average of patents per capita in 2005-2007, *rd* indicates R&D expenditures over GDP, *hk* is the population share of graduates. As control variables we include the population density and the regional share of manufacturing activities. See Appendix 2 for a detailed description of the variables.

More specifically, as a proxy of innovative activity we use the number of patents application filed at the European Patent Office (EPO) classified by priority year and by inventor's region. In case of multiple inventors, we assign a proportional fraction of each patent to the different inventors' regions of residence. Since patenting activity, especially at the regional level, is quite irregular over time we smooth the variable by computing a three-year average. Moreover, to control for the different size of the regions, the number of patents is divided by total population. Thus our dependent variable (*inn*) is measured as the yearly average of patents per million inhabitants in 2005-2007. The summary statistics, reported in Table 2, show substantial differences in patenting activity among European regions, ranging from near zero in Sud-Vest Oltenia, Romania, to 627 in the German region of Stuttgart. The high value (1.2) of the coefficient of variation (CV) confirms the great degree of spatial concentration of innovative activity which is clustered in the north-centre of Europe while little patenting activity is performed by the eastern and southern regions.

The traditional input in the KPF is R&D expenditure, rescaled for GDP, which shows an average value of 1.4. In this case, yet again, the spatial distribution in Europe is quite concentrated (CV=0.85) in Scandinavia, Central Europe (Germany, Switzerland, France) and in Southern England.

As an additional input, expected to influence the process of knowledge production at the local level, we consider the availability of human capital. Following a well-established literature we measure human capital as the share of population with tertiary education (ISCED 5-6). The spatial distribution of this variable across European regions appears more uniform (CV=0.39) and with a clearly identifiable national pattern. A high endowment of human capital characterizes the Scandinavian countries, UK, Germany, Spain while lower values are generally detected in the Eastern countries, France and Italy.

Population density is included to account for possible agglomeration effects, which especially in urban contexts are associated with more intensive innovation activity. Audretsch and Feldman (1996) emphasize that the location of manufacturing activity is one of most relevant factors that explain the spatial distribution of innovative activity, thus to control for this aspect related to the local productive pattern, we also include the regional share of manufacturing activities. The vector of control variables includes GDP per capita to account for the different levels of economic development across European regions,[13] population density to allow for possible agglomeration effects and the share of manufacturing activities to account for the regional productive pattern.[14]

Note that all the explanatory variables included in model (1) are averaged over the three-year period 2002-2004. The average values are expected to smooth away undue business cycle

[13] Following Usai (2011), the GDP per capita is included as an indicator which takes value 1 for regions with a GDP per capita level above the European average and zero otherwise. The indicator is preferred to the level of the variable since the latter induces multicollinearity problems with the other regressors, in particular with the share of R&D.

[14] Preliminary, as additional controls we have also considered social capital, employment and the settlement typology structure, which comprises both density and the existence of large urban centres; however, all of them turned out to be not significant.

effects, the lags with respect to the dependent variable are necessary to allow for a congruent time horizon for the productive inputs to unfold their effects. Moreover, lagged explanatory variables should also avoid potential endogeneity problems.

Proximity factors are included in the model in order to capture the potential role of spillover effects running along the five different dimensions suggested by the literature – geographical, institutional, technological, social and organizational.

Since the presence of spillovers induces spatial correlation in the patenting activity among the regions, the proximity factors have to be modelled accordingly. The spatial econometric literature provides two basic models to account for the existence of spillovers: the Spatial Autoregressive (SAR) model, which features a spatial regressor given by the weighted average of the all-other region response variable, and the Spatial Durbin (SD) Model, which extends the SAR model by including also the weighted average of the explanatory variables. For both specifications the weights represent the assumed interconnectivity structure among the spatial units.

It is worth remarking that, in this paper, we do not consider the Spatial Error Model (SEM), which entails spatial dependence only in the model errors, as it removes spillovers by construction. Within the SEM model, spatial dependence is not the focus of the analysis, but it is seen as a nuisance which yields non-spherical error, so it is treated just to ensure unbiased variance estimators.

Note also that we rule out the SD Model on substantive grounds, for this specification implies that the influence of neighbouring territories on the innovative performance of a certain region is mediated also by their internal inputs, i.e. R&D investments and human capital endowments, conditional on a given connectivity structure. This amounts to assuming that neighbours' R&D investments are thoroughly productive and that human capital feature a considerable degree of mobility among regions. As both these assumptions are hardly realistic in the European context,[15] we argue that it is more reasonable to envisage that innovation spillovers work through the effective level of knowledge achieved by neighbouring regions, which is proxied by the number of patent applications. Therefore, our preferred specification is the spatial autoregressive one, formalized as follows:

$$inn_i = \beta_1 rd_i + \beta_2 hk_i + \phi \; controls_i + \rho W \; inn_i + \varepsilon_i \qquad (2)$$

[15] For instance, some expenditures classified as R&D are not directly related to knowledge activities (as is the case of research laboratory buildings) and 50% of R&D is made of researchers wages, so that it is more plausible to allow for the existence of R&D spillovers in the case of the general level of production than in the case of the specific knowledge creation process. This is also confirmed by a preliminary econometric investigation based on the SD model, which resulted in not significant spillovers effects.

where W is a weight matrix, which describes the interconnectivity among regions according to one of the proximity dimensions previously discussed.[16]

In model (2), due to the presence of the spatially lagged dependent variable the interpretation of the coefficients as partial derivative no longer holds. The *total* effect on the innovation response variable caused by a unit change in one of the internal factors - either R&D or human capital - has a complex structure and can be decomposed into a *direct* and an *indirect* effect. The *direct* effect measures the change in region *i*'s dependent variable caused by a change in one of its own regressors plus a series of feedback effects (region *i* is neighbour to its neighbours so affecting them will receive in turn a feedback influence), while the *indirect* or spillover effects is due to a change in another region's regressor. It is worth noting that feedback and spillover effects occur over time through the simultaneous system of interdependence among regions, so that the effects have to be considered as the result of a new steady state equilibrium. LeSage and Pace (2009) propose summary scalar measures for direct, indirect and total effects along with their dispersion measures, which allow to draw inference on their statistical significance.

In the subsequent sections we analyse spillovers by considering one proximity dimension at a time, starting with the traditional geographical one and then following with the technological, the social and finally the organizational proximity. As will be explained in greater detail in the next section, regional institutional closeness is better dealt with by including the complete set of country dummies, so we do not propose a SAR model with the dependent variable lagged term based on the institutional structure. As the national dummies can be considered as additional control variables, the institutional kind of proximity is always included in the empirical models along with one of the other four connectivity measures.

4. Estimation strategy and model specification

In order to estimate the KPF model it is necessary to select the most adequate specification in order to properly account for the interconnectivity structure among the European regions and thus provide a more reliable estimate of the impact of both internal and external endowments of R&D and human capital on patenting activity.

---

[16] The SAR specification has recently been criticized (Partridge et al., 2012 and Gibbons and Overman, 2012) for lacking identification of the ρ parameter (see model 2) when the weight matrix is block-diagonal and idempotent. In our case the weight matrices considered do not share such properties. Moreover, Gibbons and Overman interpret ρ parameter as the causal effect of the neighbouring response variable. However, LeSage and Pace (2009) warn against such interpretation and Elhorst (2010) explains very clearly that the spillovers (indirect) effects have the complex structure of a multiplier term whose size and sign depend on both the estimated coefficients and on the weights matrix. We provide a brief description of direct and indirect effects in the main text.

As argued in section 2, innovation activity is likely to exhibit cross-regional dependence due to the emergence of a number of interactions among agents, firms and institutions located in different regions. On the basis of a preliminary investigation (see Marrocu et al, 2012), we rule out that such dependence can be considered as a nuisance yielding non-spherical errors since spatial LM tests indicate the presence of spillovers effects.[17] The occurrence of spillovers is in line with evidence on increasing returns to scale and regional disparities induced by agglomeration effects, which have been rationalized by the endogenous growth (starting with Romer, 1986) and new economic geography theories. According to the taxonomy suggested by Anselin (2003), spillovers could be local or global in nature and can be represented by basically three main spatial specifications: the spatial least squares model (SLX), features local externalities captured by the spatially lagged exogenous variables; on the contrary, the spatial autoregressive model (SAR), thanks to the spatial lag of the dependent variable, yields global spillovers. Both kinds of externalities are featured by the third encompassing model, the spatial Durbin model (SDM), which includes spatial lags for both the response and the explanatory variables. Clearly, discriminating among these different specifications can be very difficult; as argued in Anselin (2003) the problem can be approached either from an entirely empirical perspective by relying on specification tests, or on the basis of substantive theoretical arguments which should identify the nature of externalities.

We rely on both approaches. In light of the knowledge transmission mechanism described in the theoretical section, innovation externalities are believed to be more global in nature since all locations in the system are assumed to be related to each other through a multi-dimensional structure of networks and interactions. On these grounds we focus on the estimation of SDM and SAR models. However, the first specification yields implausible results with spillover effects being either not significant or with the wrong sign.[18] This result may be due to the intrinsic characteristics of the SDM specification, which entails a very complex externalities structure, and puts too strong a requirement on the data. Moreover, it is reasonable to argue that the enhancing effects on knowledge production transmitted by proximate regions, at the aggregate territorial scale (NUTS 2 regions) considered in this study, are more likely related to their innovation results rather than to their efforts to produce them.

Therefore, on the basis of these arguments, the SAR model is considered the most adequate specification, conditioned on our sample data. The related structural model implies that the effect of a given explanatory variable is the result of all the interactions among regions that have taken

[17] With respect to model (2) the LM-error test has a p-value of 0.86, while the LM-lag test exhibits a p-value of 0.02. Note that the preliminary analysis was mostly based on the geographical proximity.

[18] As a robustness test we also estimate the SLX model and the spatial Durbin error model and we find similar result for the lagged explanatory variables.

place across space and over time. Identification of the effects is obtained on the basis of the decaying behaviour of the distance or similarity measures. Thus, inference on the estimated coefficients allows to explain the interconnectivity pattern for all the locations as a function of the exogenous variables.[19]

In modelling the regional innovation activity, we adopt a specific-to-general approach, starting from a SAR specification which models the interconnectivity among regions by considering one proximity measure at a time. We begin with the proximity most commonly used in empirical studies, i.e. the geographical one, and then we proceed with the other ones. In order to account for the complementarity among them, ideally it would be preferable to specify a comprehensive model which accounts for all possible proximity factors at the same time. This would require solving an order five-multivariate optimization problem over the range of feasible values for the autoregressive parameters and thus would entail the development a new spatial econometric toolbox, which, however, goes beyond the scope of this study. As a matter of fact, in the spatial econometric literature only a variant of the spatial lag model with two spatially lagged terms for the dependent variable has been proposed so far (Lacombe, 2004). Therefore, in this study proximity complementarity is preliminarily analysed by adopting such a variant of the SAR model and deriving, in a later phase, encompassing results by resorting to model combining techniques.

Before proceeding with the detailed discussion of the results, it is worth recalling that in the case of the SAR model, the effects of the explanatory variables no longer coincide with the estimated coefficients because of the presence of the spatially lagged dependent variable; this induces feedback loops and spillover effects generated by the dependence structure of the spatial units. The *total* effect caused by a change in one explanatory variable can thus be decomposed into the *direct* effect (the change in region *i*'s dependent variable caused by a change in one of its own regressors plus the feedback effects) and the *indirect* or spillover effects (the change in region *i*'s dependent variable caused by a change in region *j*'s regressor). It is worth noting that feedback and spillover effects occur over time through the simultaneous system of interdependence among regions, so that the effects have to be considered as the result of a new steady state equilibrium. LeSage and Pace (2009) propose summary scalar measures for direct, indirect and total effects along with their dispersion measures, which allow to draw inference on their statistical significance.

[19] It is worth noting that the results obtained from the spatial specifications discussed above may be influenced by spatial heterogeneity. Differently from the case when panel data are used, which allow to treat the problem by including fixed or random cross-section and time effects, it is difficult to deal with heterogeneity when using cross-section data. However, the inclusion of three different control variables along with the complete set of national dummies is expected to alleviate the problem. The estimation of more involving models, which explicitly deal with spatial heterogeneity, such as those comprising the existence of spatial regimes or varying coefficients over space, goes beyond the scope of this paper.

5. Empirical results

5.1 Proximities and networks: a preliminary comparison

In this section we present the results of the SAR model estimated by using the proximity dimensions one at a time; this first stage analysis allows to carry out comparisons with the previous empirical literature. It is important to remark that all regressions include a set of country dummies to account for institutional closeness, such as sharing norms and a common language .

Following the extensive analysis done by Marrocu et al. (2012) the geographical matrix *G* is confined to the range 0-600 km since the spatial spillovers are localised and limited in space.[20] Similar considerations apply to the technological matrix *T*, which generates relevant spillovers only when the similarity index between the two regions is above the threshold of the 0.5 value.

Note that each proximity matrix is maximum-eigenvalue normalized; as emphasised in Keleijan and Prucha (2010), such a normalization is sufficient and avoids strong undue restrictions, as it is the case when the row-standardization method is applied. Moreover, symmetry and the importance of absolute, rather than relative, distance is maintained.[21]

The estimation results for the four KPF models based on a single proximity measure are reported in Table 3. An interesting outcome is the low variability of the estimated coefficients both for the input variables and for the controls. Considering the estimated effects in detail, the total elasticity for R&D goes from 0.20 when the regional connectivity is proxied by the social or the organizational matrix, to 0.34 in the technology proximity-based model. In the latter model the highest elasticity is also found for human capital (1.90), while the organizational-based model yields its lowest value (1.56). Thus, the first important result is that human capital is more effective than formal research expenditure in determining knowledge production at the regional level. We find that the total impact of human capital is always higher with respect to R&D in all models, ranging from a multiple of around five in the model with technological similarity to above eight in the model with social networks. As the creation of new knowledge is often based on informal learning processes and on the ability to exploit external knowledge, a well-educated labour force plays a key role in these processes. It is also worth noting that indirect effects are almost always significant and sizeable for human capital, accounting for up to 32% of the total effect in the case of the technology based model and 16% with the geography based one.

[20] The literature has emphasized the localized nature of geographical knowledge spillovers which are often limited in space (Doring and Schnellenbach, 2006). Previous findings for EU15 regions show that knowledge spillovers are confined to a range of around 300 km (Bottazzi and Peri 2003; Moreno et al. 2005), while a crucial distance of 600 km is found by Dettori et al. (2012).

[21] When the proximity weight matrix is capturing a "distance decay" type of economic behavior "scaling the rows so that the weights sum to one may result in a loss of that interpretation" (Anselin, 1988, p. 24).

Comparisons with previous similar studies on European regions, where no direct/indirect/total effects were reported, could be made only on the basis of the estimated inputs' coefficients. Our R&D estimated coefficients are very similar to the one of 0.26 reported by Moreno et al. (2005) for 17 countries, while Bottazzi and Peri (2003) present a higher value of 0.8 for 86 regions in EU12. For human capital the only two comparable studies are the one by Greunz (2003) for 153 European regions and the one by Usai (2011) for 342 regions in OECD countries, who present point estimates of 2.0 and 1.0, respectively.

As for the controls, the GDP per capita indicator and the manufacture specialization structure exhibit positive and significant coefficients across the four models indicating that knowledge creation exhibits a significant correlations with both high income levels and manufacturing productions. On the other hand, population density turns out to be not significant, plausibly because the agglomeration effects it was meant to capture are already accounted for by the GDP variable.

Turning to the coefficient of the lagged dependent variable, the first remarkable outcome is that it is always positive and statistically significant in all four models, signalling that each of the different proximity measures captures the cross-regional dependence arising from the knowledge transmission mechanisms described in section 2. More specifically, the strongest association (the spatial lag coefficient is equal to 0.31) is found for the technological proximity, which turns out to be the most important channel of knowledge spillovers, whilst geographical proximity ranks second (0.15). As far as the network dimensions are concerned, they have a relatively more modest role: the spatially lagged dependent variable exhibits a coefficient of 0.09 when it is computed using the social proximity and 0.06 in the case of the organizational one.

Comparing our results for the lagged dependent variable coefficient with previous studies, it turns out that the coefficient of the geographical proximity matrix goes from 0.09 for EU regions in Moreno et al. (2005), to 0.18 in Usai (2011) which refers to both US and EU, to a much higher value of 0.4 for the US in Carlino et al. (2007). For the technological proximity previous comparable studies are Moreno et al. (2005) with a lag coefficient equal to 0.05 and Greunz (2003) with an estimate of 0.25, who also reports that technological association is stronger than the geographical one. Our findings related to a lower effect of the social dimension confirm previous results by Maggioni et al. (2007) who found that geographical proximity has an effect that is double with respect to the relational one.

Finally, it is worth noting that the diagnostics tests (bottom panel of Table 3) are not significant across the four models, indicating that there is no remaining spatial correlation in the residuals; in other words, all the proximity measures considered are adequate candidates for capturing regional interconnectivity. This result lends further support to the SAR model, since, had the spatial functional form been misspecified, this would have shown up in the residuals yielding significant diagnostic tests.

5.2 Models with pairs of proximity matrices

As it has been remarked in the literature, the different types of proximity are expected to be complements as they represent knowledge transmission channels which reinforce each other over time and across space (Mattes, 2011). From the empirical point of view this implies that one should include all the kinds of proximity in the same estimation model. Unfortunately, the available estimation codes for spatial econometrics do not allow this first best solution and we have to look for second best procedures.

In this section we present the results for the SAR models estimated by including two different proximity-lagged terms at a time, in order to account for complementarities between pairs of knowledge spillovers channels. The two-weight matrix SAR model is specified as: $Y = X\beta + \rho_1 W_1 Y + \rho_2 W_2 Y + \varepsilon$ and it requires to solve a bivariate optimization problem over the range of feasible values for the parameters ρ_1 and ρ_2. [22] This model specification was first proposed by Lacombe (2004) to carry out a policy spending evaluation analysis within a spatial framework.[23] Such models are a useful estimation device when the connectivity among spatial units cannot be entirely captured by the traditional geographical measures (distance, contiguity, nearest-neighbours) since it also features other a-spatial kinds of links.

Results are reported in Table 4 which shows that, remarkably, most of the previously discussed results maintain their strength and significance. This is the case for the main determinants of knowledge production – R&D and human capital – the controls and the spatially lagged dependent variables. In particular, the strength of the geographical connectivity is confirmed, for all the three models where this is considered (first three columns), with an estimated average value of 0.14. The same applies for the proximity measure based on technological similarity, which exhibits a relatively higher impact (average value of 0.32) when compared with the geographical one. The regional connectivity based on both the social and the organizational proximity shows a weaker degree of dependence, with an estimated coefficient which, on average, is equal to 0.11 and 0.07, respectively. Note that when these matrices are included together (last column in Table 4) both coefficients of the spatially lagged terms are no longer significant, signalling a sort of

---

[22] See LeSage and Pace (2009) for a detailed description of the estimation procedures.
[23] We are very grateful to D.J. Lacombe for making available to us the Matlab scripts to estimate two-weight matrix SAR models.

multicollinearity problem. This is possibly due to the fact that, as remarked before, there is some overlapping in the information contained in the two matrices.

As far as the knowledge production inputs, R&D and human capital, the results provided in the previous section are broadly confirmed. The estimated coefficients are significant in all the six estimated models. In the bottom panel of Table 4 we also report the estimated direct, indirect and total effects. It turns out that human capital exhibits higher impacts, both direct and indirect, with respect to R&D, thus proving to be highly productivity-enhancing for regional innovation activities.

Finally, it is worth highlighting that spillover effects are significant for all models but only in the case of human capital. This result is consistent with the claim that R&D expenditure *per se* is not sufficient to activate knowledge externalities[24] and this, in turn, calls for policies and production devices capable of increasing the absorptive capacity of the regional systems of innovation.

Overall, the model that yields the highest total impacts is the first one, when the interdependence among regions is captured by the geographical and the technological patterns. Note that spillover effects are rather relevant, as in certain cases they are almost of the same order of magnitude as the direct ones.

5.3 Model comparison and the overall effect of knowledge spillovers

Although all the estimated models provide promising evidence on the role played by the knowledge productive inputs and on the relevance of different regional transmission channels, it is quite difficult to select a preferred model among those presented in Table 3 and 4.

Various approaches may be adopted to carry out a selection from the estimated models, some are based on testing procedures (Kelejian, 2008; Burridge and Fingleton, 2010), others on the use of information criteria or on the computation of posterior model probabilities or Bayes' factors (LeSage and Pace, 2009). In this paper, we apply the Akaike information criterion (AIC) to obtain a possible ranking of the estimated models. This has the advantage of avoiding several model comparisons, as would be the case with the testing approach. Moreover, once the "best" model, defined as the one which minimize the AIC, is found, relative probabilities of minimizing the information loss can be computed for each remaining model as a function of the difference between its own AIC value and the minimum one. A weighted multi-model could be then obtained on the basis of such probabilities (Burnham and Anderson, 2002).

[24] This is likely due to the fact that a large part of R&D expenditure is represented by researches' wages, whose effect is evidently captured by the human capital variable. Moreover, it is often the case that expenditures classified as R&D are not directly related to research activities, but rather to infrastructures and logistics, so they have basically no effect on proximate regions. A similar result has been found by Crescenzi and Rodriguez Pose (2012) for the case of the US.

The computed AIC values for the ten non-nested estimated models of Table 3 and 4 point out that the "best" model is the one based on the geographical and technological proximity (model 1 of Table 4), followed by model 4 and 5 of Table 4; the other models seem to provide relatively less support[25]. It is worth remarking that the best performing models are found among those which allow for a certain degree of complementarity among the proximity measures, and that such a complementarity turns out to be rather relevant when the technology interconnectivity is involved.

On the basis of the AIC values we thus compute the relative probabilities[26] described above in order to carry out a tentative exercise to figure out the overall spillover effects when all potential proximities are taken into account. This is, necessarily, a post-estimation computation carried out to combine the inference drawn from the four one-matrix models (Table 3) and the six two-matrix models (Table 4). The overall effects are computed analytically on the basis of the weighted average of the estimated coefficients for R&D and human capital obtained from the ten models, which are 0.21 and 1.30 respectively, and the weighted average estimates of the coefficients for the four different kinds of proximity lagged terms. For all measures, the weights are represented by the models' relative probabilities.

In order to ease the comparison of the strength of proximity dependence, the estimated coefficients of the lagged dependent variables for all combinations of matrices are summarized in Table 5, where the main diagonal reports the lag coefficient estimated in the single-proximity models, while the off-diagonal entries are the coefficients obtained from two-proximity models. The last column reports the weighted average calculated on the basis of the models' probabilities described above. We observe that, on average, dependence among regions is stronger when it is captured by the technological proximity (the average of the estimated coefficients for the technological lagged dependent variable is 0.32). The connectivity appears weaker for the geographical proximity (0.07) and the social one (0.025), while the lowest dependence is found for the organizational (0.016) proximity.

In Table 6 we report the direct, indirect and total effects computed by deriving an all-proximities multiplier for both R&D and human capital on the basis of the weighted averages of the relevant parameters. From this computational exercise, considering the calculated effects at face value, it is possible to design interesting what-if scenarios for the European regions. For example, if we conjecture an increase of the ratio between R&D expenditure and GDP of 10%, so that the actual European average value increases from 1.4% to 1.56%, this should generate a total increase

---

[25] The same ranking of the models is obtained when computing either the Bayesian or the Hannan-Quinn information criteria.

[26] The probability for model *i* are computed as: $prob_i = exp(-(AIC_i - AIC_{MIN})/2) / \Sigma_j^M exp(-(AIC_j - AIC_{MIN})/2)$, where *M* is the number of models and AIC is the bias-adjusted value of the Akaike information criterion.

of patents (per million population) of 3.3 units, from the observed average value of 105 to the new computed value of 108.3 (with 64% of the change attributable to direct effects and the remaining 36% to spillovers). On the contrary, if the 10% increase refers to human capital, entailing an increase in the share of graduates from the European average value of 10.5% to 11.6%, this would yield a total increase of 20.3 patents (per million population) generated for 13 units by regional internal efforts and for 7.3 units by knowledge spillovers, made effective by the absorption capacity of the local well-educated labour force.

We think that the computation of the all-proximity multiplier for the two KPF inputs, even with all the caveats that this kind of exercise requires, provides useful indications on the relative role and importance of R&D and human capital in determining innovation. Moreover, the finding that, in some cases, spillover effects are as large as the direct effect calls for coordination policies at regional, national and European level. These policies should recognise that regions are part of different geographical, cognitive, social, institutional and organizational structures and networks and that these dimensions require appropriate actions to favour their specific positive impacts.

6. Concluding remarks

Economists and politicians both agree that the availability of knowledge and its diffusion are crucial ingredients for fostering economic development in Europe both at the regional and national level. A similar agreement is now emerging about the idea that the diffusion of innovation depends on the relative position of each region with respect to different dimensions which go beyond the geographical space. These dimensions are mainly a-spatial and include the institutional, technological, social and organizational ones. In this paper, moving along the research line of the KPF model, we have examined these issues reaching interesting and original results on the role of internal and external factors in promoting knowledge creation at the regional level.

As far as the internal factors are concerned, we find that both R&D and human capital are essential components of technological progress, even though with quite a distinct magnitude. Once institutional proximity is considered, the latter exhibits almost six times the impact of the former. This outcome is a clear indication of the effectiveness of skilful and qualified labour force in ensuring incremental technological progress based on pervasive and continuous learning, ideas circulation and experience accumulation. This is particularly true in current economic systems where the continuous emergence of new technological trajectories calls for an encompassing and systemic capacity to understand, acquire and control original knowledge and innovations.

Regarding the external factors, we establish that all dimensions of interregional proximity and connectivity are significantly related to innovative performance, representing effective channels of knowledge transmission. Nonetheless, we find that their relative strength differs significantly. The strongest association was found for the cognitive or technological proximity: 1.5 times higher than the one based on geographical proximity and up to three-four times higher than that of social and organizational networking. The existence of a common knowledge and productive base can thus be more important than unintended interactions due to spatial proximity. Moreover, we prove that intended interactions, which model social and organizational networks, are important too, although their relevance is relatively more modest. As a consequence, we find that a sizeable part of the total effects of R&D investments and human capital endowments on the knowledge creation in a certain region derives from spillover effects coming from other regions along a composite system of interregional connections. In other words, the intensity of indirect effects vary with the proximity dimension employed, but they are all fundamental in channelling knowledge through a variety of regional interdependences.

It is worth underlining that the results associated with social and organizational proximities are likely to be driven by the inherent difficulties faced in measuring their precise content. This represents a limitation of the current study, which we plan to address in future analyses by exploiting the additional explanatory power of alternative data sources at the micro level (i.e. European social survey), which are expected to provide more reliable measures of social closeness at regional level. Another limitation, which deserves further consideration in future extensions, is related to the assessment of potential complementarities among the different proximity dimensions. The development of a comprehensive econometric framework would enable us to account for the complete range of complementarities, which are supposed to exist among the proximity dimensions, and to provide a more rigorous measure of the overall knowledge multiplier. Moreover, further research is necessary to unveil the underlying links between the aggregate regional macro level and the micro level, where individual behavior and relations are shaped along each dimension of proximity. As a matter of fact, there is strong need for micro-econometric analyses on the causal effects of industrial and regional policies, such as those by Criscuolo et al. (2012) and Antonelli and Crespi (2012), to acquire more specific indications on the more effective interventions and instruments to be implemented.

In this study we have investigated the complementary role played by five different kinds of proximity in driving knowledge transmission across the European regional innovations systems. There is by now a widespread consensus among scholars that the transfer of knowledge is significantly favoured not only by spatial closeness among agents involved in the innovation process, but also by the relations they develop within a-spatial networks, such as those shaped by institutional, technological, social and organizational links.

Nonetheless, in previous empirical literature the attention has been mostly focused on just one kind of proximity, usually the geographical one or, to a lesser extent, the technological one. At the same time several authors (Boschma 2005, Mattes 2011) have argued that, with the increasing level of economic and institutional integration within the European production context, the concurrent effect of different proximity dimensions can no longer be overlooked.

This paper contributes to the current debate by operationalizing the whole set of suggested proximity dimensions for the European regions and by analysing, for the first time, their complementary role in enhancing innovation diffusion. Our empirical analysis is carried out for a sample of 276 European regions within the prevailing KPF framework where the response variable is represented by the patents stock, while the main internal inputs are R&D investments and human capital. On the basis of our model selection strategy, a spatial autoregressive specification turns out to be the most adequate in assessing the influence of the external factors in the form of spillovers flowing along different dimensions. Beside the traditional geographical proximity, we also consider other regional interconnectivity channels represented by the institutional setting, proxied by a set of country dummies, the technological base, measured by the specialization productive structure, as well as the social and organizational networks, measured on the basis of inventors and applicant-inventor relationships occurring across different regions. The overall five-proximity multiplier for both R&D and human capital is derived on the basis of model averaging techniques applied to the various non-nested estimated models.

Four main results emerge from our empirical analysis. First, in all models considered human capital is more innovation enhancing than R&D: its total effect, which includes knowledge spillovers transmitted by proximate regions, is on average six times higher than the one associated with R&D expenditure. Second, spillover effects are significant for human capital in all models considered. This original finding indicates that it is the endowment of skilled and well-educated people which ensures that knowledge flowing from external sources can be effectively absorbed and transformed into new ideas and innovations, while high levels of R&D do not seem to grant the same desirable result. Third, all proximity dimensions considered are found to play a significant role in channelling knowledge flows. Comparing the strength of regional association captured by the different “closeness” dimensions, the technological one ranks first, followed by the geographical one; the weakest relations are found for the social and organizational networks. Fourth, we find evidence of important complementarities among the different proximities. This turns out to be rather relevant in all the cases in which the technological connectivity is involved, signalling that a common cognitive base appears to be a crucial element for conveying knowledge across regions.

Overall, the analysis presented in this paper confirms the great degree of complexity of the knowledge creation and diffusion process in the highly integrated European economic context. Our findings highlight the prominent role played by human capital in driving knowledge transfers and innovation creation and the importance of extending and strengthening regional interconnectivity along both spatial and a-spatial dimensions.

We expect the key findings summarised above to be reinforced by further investigations overcoming the limitations of the current study mainly related to the methodology and the data. Regarding the methodology, the optimal estimation strategy would require the specification of a comprehensive model which includes all proximity dimensions to fully account for their complementarities. The data limitation is mainly related to the purpose of the present study, that is to provide a general analysis of the knowledge transmission mechanisms throughout Europe, which has required the use of aggregate data since detailed information is currently available only for limited geographical contexts or for specific productive sectors.

7. Policy implications

Notwithstanding the limitations discussed above, there are some relevant and original empirical findings which allow for a better understanding of the processes of knowledge creation and diffusion in Europe. This enables us to formulate a number of policy recommendations, some with a general relevance, some others specific to the proximity dimension considered and some final ones referred to the specific situation of European neighboring countries.

The first general policy advice is that European regions still need to focus on actions aimed at increasing the endowments of well-educated labour force, given their strong and pervasive role in determining both the internal creation and the external absorption of knowledge. The impact of graduates on innovation activities is much stronger than formal R&D expenditures. New ideas, inventions, product and process innovations come mainly from the inventive capacity of well-educated people and thus education in general and universities in particular have to be central in any innovation policy.

The second general policy implication derives from the existence of several channels of interregional spillovers and externalities, which calls for a coordinated strategy able to achieve the optimal social outcome with differentiated interventions. It is increasingly clear that there is no "one size fits all" policy (Todling and Trippl, 2005) and that regions need to set different targets to

be achieved with diverse instruments. In general, policies should aim directly at investments in knowledge diffusion and absorption rather than merely investments in research and development for new ideas. Actually, this is one of the basic ideas behind the smart specialisation strategy which promotes place-based policies recently at the centre of a heated debate (see Barca, 2009 and World Bank, 2009). Thanks to such policy support, each region is expected to strengthen its competitive advantages by acquiring as much as possible from ongoing knowledge flows and, at the same time, spreading the benefits of innovation throughout the entire regional economy (Asheim et al. 2011).
Other specific policy recommendations are implied by the different spillovers transmission channels analyzed in this study. First of all, the presence of flows of knowledge which move along the technological space implies that regions should try to develop a balanced policy to create a common wide knowledge base and specific industrial platforms to maximize the absorptive capacity and its effective application. Practically, policies should support and encourage the formation of dense specialised networks among regional innovation systems, which go beyond geographical clusters. The fact that technological proximity matters even more than the geographical one in transmitting spillovers means that knowledge diffusion is facilitated within a-spatial technological clusters. This suggests the implementation of specific industrial policies to support the functioning throughout Europe of such a-spatial industrial clusters characterized by proximate technology.

Furthermore, the empirical relevance of institutional proximity implies that public coordination in the form of common procedures and standards may be crucial for avoiding opportunistic, or merely inefficient, behaviours due to lack of trust among agents in different regions. Thus, a process of effective homogenization of norms and procedures for the whole of Europe is required to help the creation of a real institutional closeness among all European areas. At the same time implementation procedures should not translate in excessive bureaucracy favouring inertia and delaying the integration with different institutional and cultural settings.

Finally, externalities arising from social and organizational interregional relations require policies designed specifically to sustain those areas where the absence or the shortage of either social or organizational capital may hamper the creation of such networks. Since these networks have an intended voluntary nature, policies have to provide a balanced set of incentives to motivate more cooperative attitudes towards economic agents located in proximate regions. Nonetheless, such inclusive policies should ensure that social relations to not happen at the detriment of market relations and competitive behaviours.

This interpretative scenario is crucial for regions in European Neighboring Countries since it enlarges the potential basin of knowledge externalities which may help their technological catching up with respect to advanced regions in the European Union. This basin, as a matter of fact, goes beyond the mere geographical basin to include the technological, the institutional, the social and the organizational ones. Such a scenario, nonetheless, has a very clear background conclusion: the hope of technological advance in poor backward regions starts from the promotion of excellence in education and skills development, and the enhancement of mechanisms underlying the diffusion of knowledge and the circulation of ideas. This will facilitate the catching up of laggard and peripheral areas and, at the same time, increase the potential innovative output of EU and ENC together within global competition. Such a goal clearly entails enhanced consistency of national and regional strategies. Strategies which recognise that each region innovation potential is unique because of different geographical, cognitive, social, institutional and organizational structures and networks, and each region requires specific local platform policies based on differentiated knowledge structures.

References

Abreu M., Grinevich V., Kitson M. and Savona M. (2008)Absorptive capacity and regional patterns of innovation, *DIUS Research Report* 08-11.

Acosta M., Coronado D., León M.D. and Martínez M.Á., (2009) Production of University Technological Knowledge in European Regions: Evidence from Patent Data, *Regional Studies*, 43, 1167-1181.

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

Anderson J.E. and van Wincoop E. (2004) Trade Costs, *Journal of Economic Literature*, 42, 691–751.

Anselin, L. (1988) *Spatial econometrics: methods and models*. Dordrecht: Kluwer.

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

Anselin, L. (2010) Thirty years of spatial econometrics, *Papers in Regional Science*, 89, 3-25.

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

Antonelli A. (2008) *Localised Technological Chage: Towards the Economics of Complexity*, London: Routledge.

Antonelli C. and Crespi F. (2012) Matthew Effects and R&D subsidies: knowledge cumulability in high-tech and low-tech industries, forthcoming in *Technological Forecasting and Social Change.*

Antonelli C., Patrucco P.P. and Quatraro F.,(2008) The governance of localized knowledge externalities, *International Review of Applied Economics*, 22, 479-498.

Archibugi D. and A. Filippetti (2011) Is the Economic Crisis Impairing Convergence in Innovation Performance across Europe? *Journal of Common Market Studies*, 49,1153-1182.

Asheim B.T., Boschma R. and Cooke P. (2011) Constructing Regional Advantage: Platform Policies Based on Related Variety and Differentiated Knowledge Bases, *Regional Studies,* 45, 893-904.

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

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

Barca, F. (2009), *An Agenda for a Reformed Cohesion Policy*, Independent report, EC, Brussels

Basile, R., Capello, R. and Caragliu, A. (2012) Technological Interdependence and Regional Growth in Europe. *Papers in Regional Science*, forthcoming, DOI: 10.1111/j.1435-5957.2012.00438.x

Boschma R. and K. Frenken (2010) The spatial evolution of innovation networks. A proximity perspective, in R. Boschma and R. Martin (eds) *Handbook of Evolutionary Economic Geography*. Cheltenham: Edward Elgar.

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

Bottazzi, L. and Peri, G. (2003) Innovation and spillovers in regions: evidence from European patent data, *European Economic Review*, 47, 687-710.

Buesa M., Heijs J. and Baumert T. (2010) The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach, *Research Policy*, 39, 722–735.

Burnham, K.P. and Anderson, D.R. (2004) *Model Selection and Inference: a practical Information-Theoretical Approach*, Springer-Verlag.

Burridge, P. and Fingleton, B. (2010) Bootstrap inference in spatial econometrics: the J test, *Spatial Economic Analysis*, 93-119.

Carlino, G.A., Chatterjee, S. and Hunt, R.M. (2007) Urban density and the rate of innovation, *Journal of Urban Economics*, 61, 389-419.

Carrincazeaux, C. and Coris, M. (2011) Proximity and Innovation, in Cooke, P., Asheim, B.T. and Boschma, R. (eds) *Handbook of Regional Innovation and Growth*. Cheltenham: Edward Elgar, pp. 269-281.

Castellacci F. (2007) Innovation and the Competitiveness of Industries: Comparing the Mainstream and the Evolutionary Approaches, *Technological Forecasting and Social Change*, 75, 1-23.

Christ J.P. (2009) New Economic Geography reloaded: localized knowledge spillovers and the geography of innovation, *FZID Discussion Papers 01-2009*, University of Hohenheim.

Coe, D.T. and Helpman, E. (1995) International R&D spillovers, *European Economic Review*, 39, 859-887.

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

Corrado, L. and Fingleton, B. (2012) Where is the economics in spatial econometrics?, *Journal of Regional Science*, 52, 210-239.

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

Crescenzi, R. and Rodríguez-Pose, A. (2012) R&D, Socio-Economic Conditions and the territorial dynamics of Innovation in the United States, forthcoming in *Growth and Change*.

Criscuolo C, Martin R., Overman H. and Van Reenen J. (2012) The Causal Effects of an Industrial Policy, *NBER Working Papers* 17842.

Dettori B., Marrocu E. and Paci R. (2012) Total factor productivity, intangible assets and spatial dependence in the European regions, *Regional Studies*, 46, 1401-1416

Doring, T. and Schnellenbach, J. (2006) What do we know about geographical knowledge spillovers and regional growth?: a survey of the literature, *Regional Studies*, 40, 375-395.

Elhorst, J.P. (2010) Applied spatial econometrics: raising the bar, *Spatial Economic Analysis*, 5, 10-28.

Engelbrecht H.J. (2002)Human capital and international knowledge spillovers in TFP growth of a sample of developing countries: an exploration of alternative approaches, *Applied Economics*, 34, 831–841.

Gertler, M.S. (2003) Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography*, 3, 75-99.

Gibbons S. and Overman H.G. (2012) Mostly pointless spatial econometrics, *Journal of Regional Science*, 52, 172-191.

Granovetter, M. (1985) Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, 91, 481-510.

Greunz, L. (2003) Geographically and technologically mediated knowledge spillovers between European regions, *Annals of Regional Science*, 37, 657-80.

Griliches Z. (1979) Issues in Assessing the Contribution of Research and Development to Productivity Growth, *Bell Journal of Economics*, 10, 92-116.

Grossman, G.M. and Helpman, E. (1990) Trade, innovation, and growth, *American Economic Review*, 80, 86-91.

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

Hollander H., Tarantola S. and Loschky A. (2009), *Regional Innovation Scoreboard (RIS) 2009*, ProInno Europe.

Iammarino S. (2005) An Evolutionary Integrated View of Regional Systems of Innovation: Concepts, Measures and Historical Perspective, *European Planning Studies*, 13, 495-517.

Jaffe A.B. (1986) Technological Opportunity and Spillovers of R&D: evidence from Firms' Patents, Profits and Market Value, *American Economic Review*, 76, 984-1001.

Jaffe A.B. (1989) Real effects of academic research, *American Economic Review*, 79, 957-70.

Kelejian H.H. and Prucha, I.R. (2010) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, 157, 53-67.

Kelejian H.H.(2008) A spatial J-test for model specification against a single or a set of non-nested alternatives, *Letters in Spatial and Resource Sciences*, 3-11.

Kirat T. and Lung, Y. (1999) Innovation and proximity - Territories as loci of collective learning processes*, European Urban and Regional Studies* 6, 27-38.

Kroll H. (2009) Spillovers and Proximity in Perspective. A Network Approach to Improving the Operationalisation of Proximity, *Fraunhofer Working Papers Firms and Region*, N. R2/2009.

Lacombe D.J. (2004) Does econometric methodology matter? An analysis of public policy using spatial econometric techniques, *Geographical Analysis*, 36, 105-118.

LeSage J.P. and Pace, R.K. (2009) *Introduction to Spatial Econometrics*. Boca Raton: CRC.

Lobo J. and Strumsky D. (2008) Metropolitan patenting, inventor agglomeration and social networks: A tale of two effects, *Journal of Urban Economics*, 63, 871-84.

Maggioni M.A., Nosvelli, M. and Uberti, T.E. (2007) Space versus networks in the geography of innovation: a European analysis, *Papers in Regional Science*, 86, 471-493.

Maggioni M.A., Uberti, T.E. and Usai, S. (2011) Treating patents as relational data: knowledge transfers and spillovers across Italian provinces, *Industry & Innovation*, 18, 39-67.

Marrocu E., Paci R. and Usai S. (2011) The complementary effects of proximity dimensions on knowledge spillovers *Working Paper CRENoS* 2011/21

Marrocu E., Paci R. and Usai S. (2012) Proximity, Networking and Knowledge Production in Europe: what lessons for innovation policy? forthcoming in *Technological Forecasting and Social Change.*

Maskell P. and Malmberg, A. (1999) The competitiveness of firms and regions. 'Ubiquitification' and the importance of localized learning, *European Urban and Regional Studies*, 6, 9-25.

Mattes J. (2011) Dimensions of Proximity and Knowledge Bases: Innovation between Spatial and Non-spatial Factors, *Regional Studies*, DOI:10.1080/00343404.2011.552493.

Miguélez E. and Moreno, R. (2011) Research networks and inventors' mobility as drivers of innovation: evidence from Europe, *Regional Studies*, DOI: 10.1080/00343404.2011.618803

Moreno R., Paci, R. and Usai, S. (2005) Spatial spillovers and innovation activity in European regions, *Environment and Planning A*, 37, 1793-1812.

Nelson R.R. and Winter, S.G. (1982) *An Evolutionary Theory of Economic Change.* Cambridge, MA: Harvard University Press.

O´hUallacha´in B. and Leslie T. (2007) Rethinking the regional knowledge production function, *Journal of Economic Geography*, 7, 737-752.

Oerlemans L. and Meeus M. (2005) Do organizational and spatial proximity impact firm performance? *Regional Studies,* 39, 89–104.

Paci, R. and Usai, S. (2009) Knowledge flows across the European regions, *Annals of Regional Science*, 43, 669-690.

Parent O. and LeSage J.P. (2012) Determinants of knowledge production and their effect on regional economic growth, *Journal of Regional Science*, 52, 256-284.

Partridge M.D., Boarnet M., Brakman S. and Ottaviano G. (2012) Introduction to the Symposium on Spatial Econometrics: Whiter spatial econometrics?, *Journal of Regional Science*, 52, 167-171.

Picci L. (2010) The internationalization of inventive activity: A gravity model using patent data, *Research Policy,* 39, 1070-1081.

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.

Rallet A. and Torre A. (1999) Is geographical proximity necessary in the innovation networks in the era of the global economy?, *GeoJournal*, 49, 373–380.

Rodríguez-Pose, A. and Crescenzi, R. (2008) Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe, *Regional Studies*, 42, 51-67.

Romer, P.M. (1986) Increasing Returns and Long-run Growth. *Journal of Political Economy*, 94, 1002-1037.

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

Tappeiner G., Hauser C. and Walde J. (2008) Regional knowledge spillovers: Fact or artifact?, *Research Policy*, 37, 861–874.

Todling F. and Trippl M. (2005) One size fits all? Towards a differentiated regional innovation policy approach, *Research Policy* 34, 1203–1219.

Torre A. and Gilly J.P. (2000) On the analytical dimension of proximity dynamics, *Regional Studies*, 34, 169–180.

Usai S. (2011) The geography of inventive activity in OECD regions, *Regional Studies*, 45, 711-731.

World Bank (2009) *Reshaping Economic Geography: World Development Report*, The World Bank, Washington, DC, 6 November.

Appendix 1. Regions and NUTS level

Code	Country	NUTS	Regions
AT	Austria	2	9
BE	Belgium	2	11
BG	Bulgaria	2	6
CH	Switzerland	2	7
CY	Cyprus	0	1
CZ	Czech Republic	2	8
DE	Germany	2	39
DK	Denmark	2	5
EE	Estonia	0	1
ES	Spain (a)	2	16
FI	Finland	2	5
FR	France (a)	2	22
GR	Greece	2	13
HU	Hungary	2	7
IE	Ireland	2	2
IT	Italy	2	21
LT	Lithuania	0	1
LU	Luxembourg	0	1
LV	Latvia	0	1
MT	Malta	0	1
NL	Netherlands	2	12
NO	Norway	2	7
PL	Poland	2	16
PT	Portugal (a)	2	5
RO	Romania	2	8
SE	Sweden	2	8
SI	Slovenia	2	2
SK	Slovakia	2	4
UK	United Kingdom	2	37

(a) Territories outside Europe are not considered

Appendix 2. Data sources and definition for variables and proximity matrices

Variable		Primary Source	Years	Definition
Patent	INN	EPO	average 2005-2007	total patents published at EPO, per million population
Research & Development	RD	Eurostat	average 2002-2004	total intramural R&D expenditure, over GDP
Human Capital	HK	Eurostat	average 2002-2004	population aged 15 and over with tertiary education (ISCED 5-6), over total population
Population density	DEN	Eurostat	average 2002-2004	Population per km^2, thousands
GDP per capita	GDP	Eurostat	average 2002-2004	value 1 for regions above the average and zero otherwise
Manufacture specialisation	MAN	Eurostat	average 2002-2004	manufacturing employment over total employment
Settlement Structure Typology	SST	ESPON project 3.1 BBR	1999	1=less densely populated without centres, 2=less densely populated with centres, 3=densely populated without large centres, 4=less densely populated with large centres, 5= densely populated with large centres, 6=very densely populated with large centres

Proximity matrix		Primary Source	Years	Definition
Geographical	G	own calculation		inverse of distance in Km
Institutional	I	own calculation		binary matrix: value 1 if the two regions belong to the same country and 0 otherwise
Technological (patent)	T	OCSE Pat-Reg	average 2002-2004	similarity index based on 2-digit SIC 44 sectoral shares of patenting activity
Technological (employment)	Te	Eurostat, Structural Business Statistics	1999	similarity index based on 17 manufacture and knowledge intensive sectoral shares of employment
Social	S	OCSE Pat-Reg	average 2002-2004	co-inventorship relation among multiple inventors of the same patent by inventors' region (intra regions relationships are not considered)
Organisational	O	OCSE Pat-Reg	average 2002-2004	applicant-inventors relation of the same patent by region of residence (intra regions relationships are not considered)

Table 1. Summary statistics for proximity matrices

Proximity matrices	Units of measurement	Min	Max	Mean	Var. coeff.	Links % *
Geographical	km	17.86	4574.57	1370.15	0.56	-
Technological	index [0 , 1]	0.05	0.94	0.70	0.18	-
Social	num links	0.00	137.84	0.16	10.68	18.18
Organisational	num links	0.00	480.13	0.58	10.52	17.11

* % of total cells, excluding the principal diagonal

Sample correlation coefficients			
	Geographical	Technological	Social
Technological	0.200		
Social	0.120	0.070	
Organisational	0.113	0.069	0.740

Table 2. Summary statistics for dependent and exogenous variables

Variable	Unit of measurement	Min	Max	Mean	Var. coeff.
Patent	per million pop	0.20	627.6	105.4	1.20
Research & Development	over GDP, %	0.07	7.6	1.4	0.85
Human Capital	over total population, %	3.51	23.3	10.5	0.39
Population density	thousands per km^2	3.08	9049.6	331.3	2.47
Manufacture specialisation	over total empl., %	3.67	36.2	17.3	0.37
GDP per capita	value 1 for regions above the average and zero	0.00	1.00	0.56	0.89

Table 3. KPF with different proximity measures: Geographical (G), Technological (T), Social (S), Organisational (O) and Insitutional (I)

Dependent variable: Patents, 2005-2007 average per capita values

Estimation method: SAR

	1	2	3	4
Proximity matrix:	G	T	S	O
Production inputs				
R&D	0.233 **	0.226 **	0.180 *	0.190 *
	(2.234)	(2.282)	(1.754)	(1.863)
Human capital	1.520 ***	1.294 ***	1.491 ***	1.459 ***
	(5.049)	(4.282)	(4.944)	(4.826)
Control variables				
GDP per capita indicator	0.430 ***	0.578 ***	0.484 ***	0.501 ***
	(2.542)	(3.614)	(2.929)	(3.045)
Manufacture specialisation	0.840 ***	0.817 ***	0.925 ***	0.946 ***
	(4.784)	(4.779)	(5.449)	(5.574)
Population density	0.043	0.061	0.058	0.059
	(0.651)	(0.975)	(0.898)	(0.908)
Institutional proximity (I)				
country dummies	yes	yes	yes	yes
Spatial lag (ρ)	0.149 **	0.311 ***	0.088 **	0.056 *
	(2.223)	(3.508)	(1.937)	(1.725)
Adj-R^2	0.812	0.816	0.811	0.810
Effects estimates (a)				
R&D				
direct	0.230 **	0.229 **	0.179 *	0.186 *
indirect	0.043	0.108	0.017	0.011
total	0.273 **	0.337 **	0.196 *	0.197 *
Human capital				
direct	1.529 ***	1.293 ***	1.488 ***	1.466 ***
indirect	0.285 *	0.602 **	0.148 *	0.089
total	1.814 ***	1.895 ***	1.636 ***	1.555 ***
Diagnostics				
LM error test for SAR model residuals	0.140	0.009	0.119	0.001
p-value	0.708	0.924	0.730	0.992

Observations: 276 regions

All variables are log-transformed

For all the explanatory variables the values are averages over the period 2002-2004

All proximity matrices are max-eigenvalue normalized

Asymptotic t-statistics in parenthesis; significance: *** 1%; ** 5%; * 10%

(a) We report only the effects for the main interest explanatory variables

Table 4. KPF with two weight matrix models

Dependent variable: Patents, 2005-2007 average per capita values

Estimation method: SAR

	1	2	3	4	5	6
Proximity matrices included	G, T	G, S	G, O	T, S	T, O	S, O
Production inputs						
R&D	0.239 **	0.199 **	0.206 **	0.180 *	0.189 **	0.180 *
	(2.439)	(1.999)	(2.067)	(1.835)	(1.923)	(1.795)
Human capital	1.325 ***	1.519 ***	1.494 ***	1.290 ***	1.247 ***	1.481 ***
	(4.509)	(5.070)	(4.987)	(4.380)	(4.233)	(4.913)
Control variables						
GDP per capita indicator	0.455 ***	0.397 **	0.400 **	0.509 ***	0.524 ***	0.485 ***
	(2.880)	(2.465)	(2.486)	(3.214)	(3.309)	(2.990)
Manufacture specialisation	0.704 ***	0.839 ***	0.850 ***	0.791 ***	0.811 ***	0.931 ***
	(4.259)	(4.984)	(5.044)	(4.777)	(4.899)	(5.490)
Population density	0.02	0.03	0.027	0.034	0.032	0.057
	(0.315)	(0.469)	(0.420)	(0.545)	(0.518)	(0.891)
Institutional proximity (I)						
country dummies	yes	yes	yes	yes	yes	yes
Spatial lag - 1st proximity matrix	0.158 **	0.129 *	0.138 **	0.324 ***	0.331 ***	0.067
	(2.413)	(1.880)	(2.038)	(3.640)	(3.708)	(0.835)
Spatial lag - 2nd proximity matrix	0.318 ***	0.069	0.048	0.099 **	0.069 **	0.018
	(3.592)	(1.763)	(1.467)	(2.184)	(2.156)	(0.316)
Adj-R^2	0.82	0.813	0.813	0.819	0.819	0.811
Estimated effects [(a)]						
R&D						
direct	0.239 **	0.199 **	0.208 **	0.181 *	0.189 *	0.180 *
indirect	0.244	0.050	0.049	0.144	0.133	0.017
total	0.483 **	0.249 **	0.257 **	0.325 *	0.322 *	0.197 *
Human capital						
direct	1.325 ***	1.519 ***	1.502 ***	1.301 ***	1.259 ***	1.495 ***
indirect	1.344 *	0.383 **	0.350 **	1.028 **	0.922 **	0.143
total	2.669 ***	1.902 ***	1.852 ***	2.329 ***	2.181 ***	1.638 ***

Observations: 276 regions

All variables are log-transformed

For all the explanatory variables the values are averages over the period 2002-2004

All proximity matrices are max-eigenvalue normalized; G=geographical (0-600 km), T=technological (index>0.5), S=social and O=organisational

Asymptotic t-statistics in parenthesis; significance: *** 1%; ** 5%; * 10%

[(a)] We report only the effects for the main interest explanatory variables

Table 5. Comparing estimated coefficients of spatial lags for different proximities measures

		Second proximity matrix included				**Weighted**
Proximity matrix considered		G	T	S	O	**average**[a]
Geographical proximity	G	0.149	0.158	0.129	0.138	0.067
Technological proximity	T	0.318	0.311	0.324	0.331	0.319
Social proximity	S	0.069	0.099	0.088	0.067	0.025
Organisational proximity	O	0.048	0.069	0.018	0.056	0.016

Diagonal entries are the estimated rho coefficients of the Table 2 one-weight matrix SAR models
Off-diagonal entries are the estimated rho coefficients of the Table 3 two-weight matrix SAR models
All the regressions include also the institutional proximity measured by the country dummies
[a] Weights are given by model probabilities obtained on the basis of AIC values.

Table 6. Combined effects of the KPF inputs

Dependent variable : Patents, 2005-2007 average per capita values

	weighted average
R&D	
direct	0.211
indirect	0.119
total	0.329
Human capital	
direct	1.298
indirect	0.730
total	2.028

Effects are computed on the basis of the weighted averages for the inputs coefficients and the spatial autocorrelation coefficients. Weights are given by model probabilities obtained on the basis of AIC values

AQR UB IREA

Utrecht University

SEED CENTER
SOUTH & EAST EUROPEAN DEVELOPMENT CENTER

CRENoS

LSE

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Brunel UNIVERSITY LONDON

Université Jean Monnet SAINT-ETIENNE

UNIVERSITAS QUINQUEECCLESIENSIS 1367

Leibniz Universität Hannover

UNIVERSITY OF TARTU

STATE UNIVERSITY HIGHER SCHOOL OF ECONOMICS

Cadi Ayyad University

ICBSS INTERNATIONAL CENTRE FOR BLACK SEA STUDIES BSEC

IEMed. Institut Europeu de la Mediterrània

האוניברסיטה העברית בירושלים
The Hebrew University of Jerusalem

TÜBİTAK