Proximities and the Intensity of Scientific Relations: Synergies and Non-linearities

Roberta Capello*, Andrea Caragliu*,§

* Politecnico di Milano
Building Environment Science and Technology
Piazza Leonardo da Vinci, 32
20133 Milan (MI)
Italy
Email address: roberta.capello@polimi.it, andrea.caragliu@polimi.it

§ Corresponding author

Abstract

Several notions of proximity affect the intensity of scientific relations with different strength. Insufficient attention has been paid to the interrelations between such forms of proximity, each one assumed to hamper the flow of goods, ideas, and spillovers on its own, but not in relation to each other. This paper aims to fill this gap. This goal is pursued with the use of a new and rich data base with data for all NUTS2 regions of the EU27 countries on many forms of proximities (including geographic, social, cognitive, and specialization proximity), and on their interaction with spatial proximity. The measures between couples of regions along these dimensions are used to explain the intensity of the real scientific cooperation among the 264 European regions, measured with the number of collaborations in European Framework Programme 5 projects, following the road paved by previous works.

Keywords: Specialization Proximity, Social Proximity, Cognitive proximity, Scientific Collaboration.

1. Introduction[1]

For decades now space has been considered in various forms in the economics literature. Since the seminal contributions by Christaller, Lösch, and Isard, space (and in particular geography) has been considered as a major factor hampering, or, at times, facilitating, economic interactions.

More recently, some additional reflections emerge on the role of spatial proximity. Firstly, geographical proximity may actually hide other types of similarities among agents in the social, technological, and cognitive sphere; spatial proximity is in this sense a proxy for other kinds of proximities, like social and cognitive proximities, that are higher in compact geographical areas. Secondly, the space where economic interactions take place is much more complex than that summarized by pure geography. In fact, modern theories emerged about the role of relational and institutional space in fostering the innovative behavior of regions. As a consequence, theories have been extended to encompass different ways to conceive of space in the analysis of economic interactions.

A vast literature has already been developed on this theme, with the goal to empirically highlight the role of complex forms of proximities on scientific cooperation, enriching the oversimplifying assumption that physical proximity – via epidemic contacts – enhances higher knowledge spillovers.

[1] A first version of the paper has been presented at the International Tinbergen Institute Workshop 2012 and at the 9th World Congress of the Regional Science Association International.

There is still room for further research on this topic. In fact, so far empirical analyses have failed to simultaneously account for all types of proximities potentially relevant for the diffusion of knowledge. Besides, and more importantly, insufficient attention has been paid to the mutual influences that different types of proximities may exert. In particular, the effects of physical proximity on the exchange of scientific knowledge may be influenced by other types of proximity, viz. cognitive, social and specialization proximity; in other words, synergies may exist between geographical and non-geographical proximity measures. Thirdly, and lastly, the relationship between scientific cooperation and the different kinds of proximities may be subject to increasing or decreasing returns; the existence of non-linearities is thus a relevant, and insufficiently studied, aspect in the field of scientific cooperation.

This empirical paper enters the debate on proximities and scientific cooperation, with the aim to fill the above mentioned gaps. A novelty of the paper is also the comprehensive way in which different types of proximity are measured. Most importantly, we build a novel measure of 'inter-regional cognitive proximity', by extending Boschma's concept of cognitive proximity to an inter-regional setting; moreover, we identify specialization and social proximities as similarities between pairs of regions in terms of the regions' social and specialization profiles. Each proximity has a clear definition, logically followed by a quantitative indicator.

The paper is structured as follows. In Section 2 the literature needed to correctly position this contribution is critically summarized; the Section concludes with the three research questions this paper answers to. Section 3 offers a description of the empirical strategy and the methodology to demonstrate the paper's research questions. In Section 4 the data set assembled for this paper and the indicators built to capture non-geographical forms of proximity are described. Section 5 discusses the main empirical findings. Finally, Section 6 concludes.

2. Literature review

2.1 Proximity and knowledge spillovers measurement

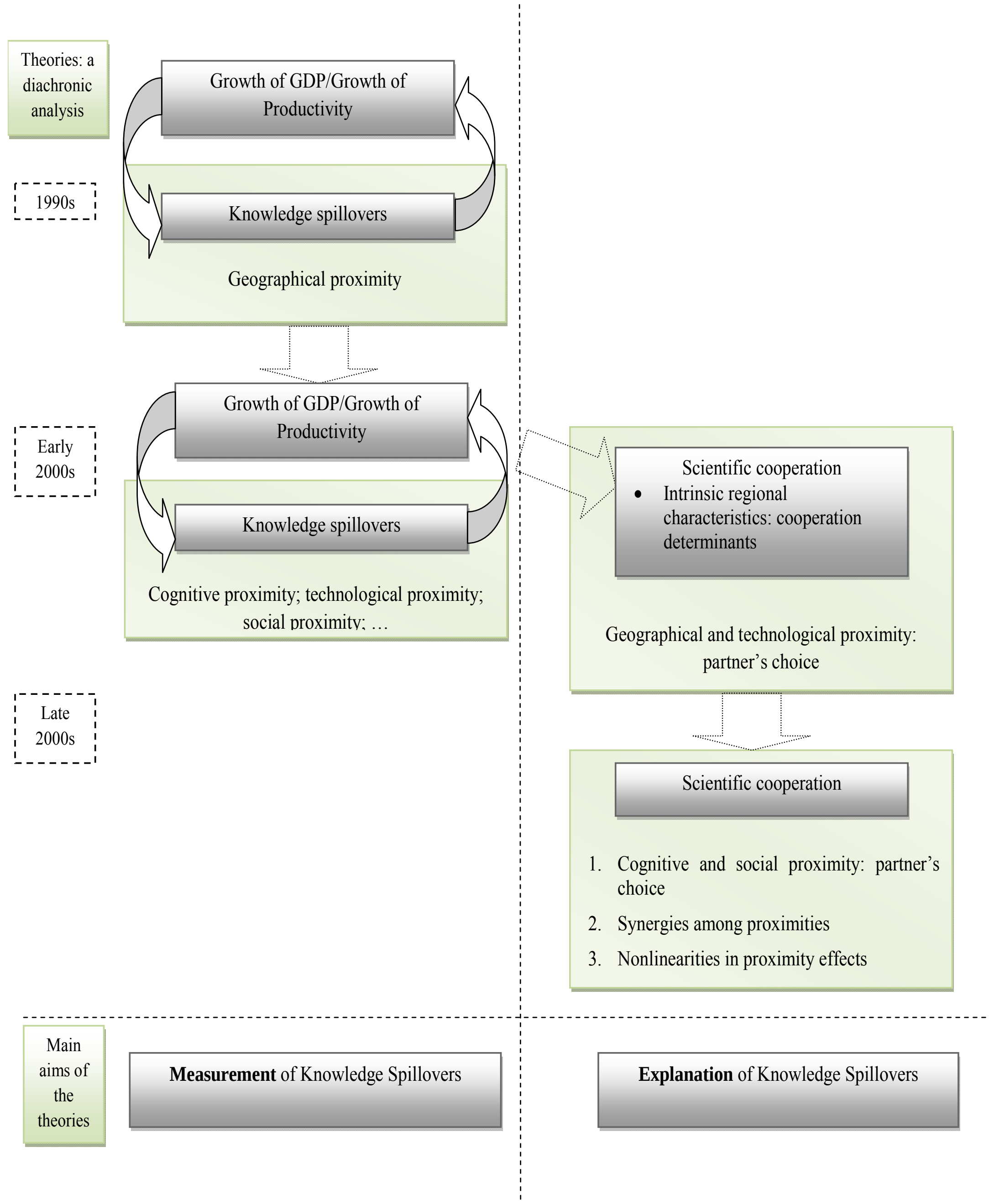
Around the mid 1980s, the economic growth literature witnessed the emergence of theoretical and empirical evidence on the fact that the productivity level of a country, or region, is affected not only by the extent to which local firms invest in R&D activities, but also by the potential access to external R&D stocks (Coe and Helpman, 1995). While initially based on pure geographical space as means of knowledge diffusion, empirical studies have in the last decade shed some light on the main channels of knowledge diffusion. Among such vehicles, much evidence has been identified for Foreign Direct Investment (e.g. van Pottelsberghe de la Potterie and Lichtenberg, 2001); in fact, through reverse engineering, firms can acquire technology embedded in traded goods (MacGarvie 2005; Padilla-Pérez 2008). However, a major role in this respect has been, and is still, played by geographical distance.

In fact, if knowledge is a partially public, partially non-rival good, it can at least partially diffuse over space. According to an epidemic viewpoint, knowledge diffusion decreases with geographical distance, because the probability of running into useful knowledge decreases as the distance between a Knowledge-Generating Institution and the knowledge users increases. The idea that geographical distance hampers knowledge flows represents the rationale for the 1990s wave of studies on knowledge spillovers (Figure 1).

Until the early 1990s the lack of appropriate measures of knowledge flows and the insufficient computing power available to scholars made the empirical validation of the above mentioned theories complicated. Simultaneous effort by industrial and regional economics lead to the emergence of two techniques to assess knowledge spillovers. Pioneering work by Jaffe and coauthors (Jaffe, 1986, 1989) identified knowledge spillovers as potential flows of knowledge originating from R&D-generating institutions within a Knowledge Production Function

(henceforth, KPF) framework. Subsequently, the KPF approach was perfected with the improvement of spatial econometric techniques (Anselin, 1988). Spatial econometrics offered a convincing toolbox to quantify the spatial impedance offered by geographical features (and in particular distance) in knowledge diffusion processes; these techniques have been applied to both the KPF and to standard growth models extended to include the effects of spatially-lagged productivity growth on local performance (Ertur and Koch, 2007). Around the same years, the idea that knowledge does leave a paper trail in patent citations led to the patent citations literature (Jaffe et al., 1993; Thompson, 2006), which provided a concrete measure of knowledge spillovers, although limited to a specific type of knowledge, viz. technological contents embedded in inventions and manufactured products (Figure 1).

Figure 1. The evolution of the literature on proximity and knowledge spillovers and the main aims of the theories



The spatial econometric literature is based on a set of simplifying assumptions, which help in the estimation process but reduce the interpretative power of these analyses. On the one hand, knowledge spillovers are only implicitly measured, by assuming that the geographical nearness

or the technological interdependence once again driven by geographical proximity suffice to determine knowledge diffusion. On the other hand, most such studies model technological interdependence as taking place mostly in geographical space. However, as correctly pointed out, "*countries may be considered as located in some general socio-economic and institutional or political space, defined by a range of factors. Implementation of spatial methods thus requires accurate identification of their localisation in such a general space. Ideally, such a matrix should be theory-based*" (Ertur and Koch, 2011, p. 236). If this statement is true for countries, it is even more binding for regions, where identity and the sense of belonging are most developed.

The second abovementioned critique (viz., the validity of geographical space as the only environment where knowledge diffuses) has been questioned both from a theoretical as well as from an empirical perspective, in particular calling for paying more attention to the cognitive dimension being implicitly assumed behind geographical proximity in the explanation of knowledge spillovers. In fact, cognitive, social, and specialization proximity represent the most relevant difference between a simple cluster and an industrial district/innovative milieu (Boschma, 2005; Capello, 2009; Breschi and Lissoni, 2001[2]). Critical issues motivated by the absence of a convincing explanation for knowledge diffusion channels have been subsequently verified empirically through spatial econometric studies where the definition of proximity underlying the construction of knowledge spillovers measures have been variously defined (Figure 1). Although the use of non-geographical proximities in modeling knowledge spillovers has been sometimes criticized, "*Proximity in geographical, industrial, and technical space matters here in that it provides reluctant and sceptic, risk-adverse adopters the opportunity to assess the actual profitability of the new technology and hence to adopt it*" (Antonelli 2003, pp. 9-10).

Empirical verifications of the role of non-geographical proximities in shaping knowledge diffusion include, among others, Autant-Bernard (2001) and Autant-Bernard and LeSage (2011), which dealt with technological (i.e. specialization) proximity; Spolaore and Wacziarg (2009), which bring genetic distance in shaping income differentials to the fore; Agrawal et al. (2008), which focuses on social proximity; and Basile et al. (2012), which deals with relational, social, and specialization proximity.

2.2 Proximity and knowledge spillovers determinants

In the studies above summarized, knowledge spillovers are mostly measured as potential flows, with the notable exceptions of works based on patent citations. In fact, spatial econometric models assume that the potential access to a wealth of external knowledge, mediated by one or more forms of proximity, represents a suitable proxy for real knowledge flows.

In order to complement such analyses, and partially solve the above-mentioned limitation, a more recent literature dealt with the determinants of knowledge spillovers and knowledge flows (Figure 1). Most such, mostly empirical, studies deals with the determinants of scientific cooperation, proxied in different manners. Along with the traditional spatial impedance offered by pure geographical space, as the previous literature on knowledge spillovers, these works relate the extent to which knowledge-generating institutions (such as research centres, individual inventors, and universities) cooperate for scientific purposes to originating institutions' characteristics, as well as to other forms of proximity. The aim of the knowledge spillovers literature moves therefore from the measurement of the effects of knowledge spillovers towards the explanation and interpretation of their existence.

[2] The need to address the externalities stemming from individual decisions at the social level, thereby suggesting the rational for the effects of social distance on economic interactions, is also at the core of Akerlof's work. See for instance Akerlof (1997).

Among the most influential studies in this sense one can include Maggioni and Uberti (2007) and Maggioni et al. (2009), who focus on the role played by non-spatial, hierarchical networks of cooperation in knowledge production; Mora and Moreno (2010), which demonstrates that specialization proximity complements geographical nearness in fostering knowledge flows across space; Autant-Bernard and LeSage (2011), which analyzes the extent of inter- and intra-industry knowledge spillovers channeled by industrial proximity; and Frenken et al. (2010) and Ponds et al. (2007), focusing on the synergic nature of geographical and relational proximity in research collaborations.

Theoretically, the choice of collaborating for scientific purposes has been modeled as a bilateral decision based on cognitive, relational, and structural embeddedness (Cowan et al., 2007). Recent empirical work demonstrating this theoretical assumption includes Autant-Bernard et al. (2007), which relate scientific cooperation to the position of individuals within an international network (and therefore deal with social distance); Scherngell and Barber (2009, 2011), whose main result is the assessment of the differential impacts of geographical and specialization proximities (with the addition of institutional factors in their 2011 contribution) on scientific collaborations, with the dependent variable measuring scientific collaborations, as in the present paper, measured by co-participations to EU Framework Programme 5 projects.

Simultaneously, around the mid 2000s, a relevant literature originated by the pivotal contribution by Jan Tinbergen on the determinants of trade flows (Tinbergen, 1962) included non-geographical characteristics of partners in the explanation of trade flows (Groot et al., 2004), FDIs (Lankhuizen et al., 2011), and, more recently, migrations (Belot and Ederveen, forthcoming).

Much has been done so far. However, we believe there is still room for further reflections, since limitations in the existing literature can still be identified, and possible solutions to such shortcomings can be suggested, namely:

1. Previous studies mostly included one or two additional proximity effects in the evaluation of the determinants of scientific collaboration. However, since geographical proximity is at most a good proxy for the underlying real proximity relations, omitting one or more relevant proximity effects may in fact offer biased estimates of the real impact of geographical proximity on scientific collaboration.[3]

   In this paper, we extend the notions of proximity accounted for, by relating scientific collaboration to cognitive, specialization, and social distance. In this work, such **proximity effects are jointly measured**.

2. Relatively neglected in previous analysis is our understanding of whether simple geographical proximity is enough to explain research collaboration, or if instead spatial proximity plays a more relevant role in selecting scientific partners when they are also cognitively, specialization-wise, and socially compatible.

   In this paper, we investigate possible **synergic effects between geographical and non-geographical proximity measures**. In particular, we verify the extent to which cognitive, specialization, and social proximity interact with geographical distance in fostering scientific cooperation.

3. Previous work on the determinants of research collaboration has consistently assumed that the effects of non-geographical proximity obeys a linear or exponential law, which, by definition, has no maximum or minimum. However, a theoretical rationale has been suggested for the need to verify the possible existence of an "optimal" proximity, defined over different concepts and with the underlying goal to maximize some innovative

[3] This point is in fact demonstrated in our empirical estimates. See Section 5 for further details.

process: "*Not only too little, but also too much proximity may be detrimental to interactive learning and innovation*" (Boschma, 2005, p. 61).

In this paper, we inspect the existence of **nonlinear effects in the impacts of non-geographical proximities on scientific collaboration**.

These three contributions lead to the following three research questions for this paper:

RQ1 *What is the joint effect of geographical distance and non-geographical proximity on scientific collaboration?*

RQ2 *Does geographical distance synergically interact with cognitive, specialization, and social proximity in fostering scientific collaboration?*

RQ3 *Do nonlinearities exist in the effects of proximity on scientific collaboration?*

3. The empirical model

We measure scientific collaboration with bilateral scientific collaborations in EU Framework Programme 5 (FP5) projects. Previous works in this line of research identified two main determinants of scientific collaboration:

- internal (to the region/institution) scientific collaboration-enhancing factors;
- external characteristics (bilateral, i.e. related to different types of close-ness between pairs of regions).

In this paper we bring forward this research and provide a comprehensive analysis of the different impacts of both internal, as well as external, scientific collaboration-enhancing factors.

Following Scherngell and Barber (2009), we model cross-regional research collaborations as a function of a set of distance-decay effects, or factors of spatial impedance, as well as of region-specific specific research determinants.

We define regions as either i or j , $i \neq j$, i= 1,..., n, j= 1,..., n. This implies that we do not evaluate intra-regional collaboration (i.e. FP5 projects where more than a single Knowledge-Generating Institution belonging to a certain region took part), and we observe the full set of scientific collaborations in the nXn matrix of possible cross-regional research partnerships.

Let p_{ij} be the number of projects to which Knowledge-Generating Institutions of regions i and j took part in the period 1998-2002. All p_{ij} represent the elements in matrix P, i.e. the probability matrix of all possible pairs of region where scientific collaboration took place.

The basic research collaboration model can be formalized as:

$$P_{ij} = X_{ij} + \varepsilon_{ij}\,, \tag{1}$$

with $i \neq j$, where:

$$X_{ij} = R\&D_{ij} \cdot PROX_{ij}\,. \tag{2}$$

Equations (1) and (2) state that the probability P_{ij} that two regions cooperate for scientific purposes depends on internal (to the region) scientific determinants, but also on the various forms of friction separating pairs of regions (*PROX*). Formally, we require $E\left[\varepsilon_{ij}\middle|p_{ij}\right]=0$. In other words, in each pair of regions the probability of collaboration should depend only on the factors (internal scientific determinants and external/bilateral spatial friction) here described; on average, deviations from such model should be equal to 0.

The structure of this model is clearly based on a gravitational structure, where the intensity of interactions between pairs of regions depends on characteristics inherent to each region as well as on various concepts of distance between elements of the regions' pair.

Eq. (2) is in implicit form. It is standard in most similar studies to (often implicitly) assume a Cobb-Douglas specification. This specification is, in fact, more tractable than most others, and allows to avoid implausible assumptions about the elasticity of the function arguments (Uzawa, 1962). Such choice allows in addition the log-linearization of the main equation, thereby identifying elasticities when empirically estimating the model.

Besides, in this paper the notions of proximity accounting for are extended. In fact, we tackle the role of cognitive, specialization, and social proximity, along with the traditional geographical dimension, as a factor enhancing research collaboration. In addition, we also verify whether regions characterized by relevant differentials in R&D investment and product innovation intensity also tend to cooperate more.

Including both the formalization of eq. (2) as a Cobb-Douglas specification, and the abovementioned notions of proximity, we can state the following explicit function:

$$P_{ij} = \begin{pmatrix} SIZE_{ij}^{\alpha} \cdot GEO\_DIST^{\beta_0} \cdot COG\_PROX_{ij}^{\beta_1} \cdot SPEC\_PROX_{ij}^{\beta_2} \cdot SOC\_PROX_{ij}^{\beta_3} \cdot R\&D\_DIFF_{ij}^{\beta_4} \cdot \\ \cdot PRODINN\_DIFF_{ij}^{\beta_5} \end{pmatrix} + \varepsilon_{ij}, \tag{3}$$

where the variables refer respectively to the size of the analyzed regions (both in terms of knowledge size, and of physical size), the geographical distance, the cognitive, specialization, and social proximities, and finally structural differences in R&D and innovation activity difference between regions, respectively.

Finally, eq. (3) can be log-linearized to obtain

$$\begin{aligned} p_{ij} = \alpha_0 + \alpha_0 size_{ij} + \beta_0 geo\_dist_{ij} + \beta_1 cog\_prox_{ij} + \beta_2 tech\_prox_{ij} + \beta_3 soc\_prox_{ij} + \beta_4 r\&d\_diff_{ij} + \\ + \beta_5 prodinn\_diff_{ij} + \varepsilon_{ij} \end{aligned}, \tag{4}$$

where lowercase letters refer to the log-linearized variables. Equation (4) is our baseline model. Section 4 presents now the indicators used to measure the complex set of proximity/distance measures captured in this analysis.

4. Indicators

The data set comprises several distance matrices calculated on the basis of proximity concepts summarized in Section 2.

a) Scientific cooperation

Following the road paved by previous works on the topic of scientific collaboration (e.g. Scherngell and Barber, 2009, 2011; Autant-Bernard et al., 2007), we measure scientific collaboration with bilateral scientific collaborations in EU Framework Programme 5 (FP5) projects, actually measuring the ex-post intensity of cooperation.

The indicator of scientific cooperation has been calculated by the authors as the count of Framework Research Programmes to which institutions belonging to each of the 69,696 (= 264^2) possible pairs of NUTS2 regions jointly participated. The matrix has been next triangularized (i.e. each co-participation has been counted just once), and the main diagonal, bearing the FP5 projects where institutions of the same NUTS2 region co-participated, eliminated. This implies a total number of possible observations equal to (69,696/2)-264, i.e. 34,584.

b) Geographical distance

The definition of geographical distance is based on distance between NUTS2 centroids in kilometers. Therefore, each entry d_{ij} is based on the following formula:

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \tag{5}$$

Following the recent theoretical call for an extension of the notions of space beyond pure geography (see Section 2), the use of geographical distance as a factor hampering scientific collaboration between regions should actually capture other underlying effects, related to other forms of relatedness between pairs of regions. Econometrically, this could be demonstrated by showing that after including other forms of proximity, the absolute relevance of geographical proximity abates. This hypothesis will be tested later in this paper (Section 5).

c) Cognitive proximity

A recent approach to proximity posited by evolutionary economic geography analysed in depth the two main components of knowledge base diversity externalities, i.e., those externalities accruing to regional growth and stemming from a diversified knowledge composition. In particular, such externalities have been classified into those stemming from cognitive distance (i.e., those arising from pure knowledge diversity), and those originating from cognitive proximity, viz. those externalities connected to the presence within an area of knowledge neither too distant, nor too close (Boschma, 2005). The idea, following Noteboom (2000), is that in order to trigger learning mechanisms, industries must enjoy a certain degree of knowledge distance in order to have something to learn, while at the same time being technologically compatible. In fact, "*it is unclear what a pig farmer can learn from a microchip company even though they are neighbors*" (Boschma and Iammarino, 2009, p. 292).

So far, the concept of cognitive proximity (based on related variety) has been operationalized at the regional level by calculating indicators of specialization relatedness across sectors. This concept has been mostly proxied by means of entropy measures (e.g. in Frenken et al., 2009) on the basis of employment or value added data at various levels of sectoral disaggregation. In particular, entropy would be best captured by elaborating indicators which correlate positively with sectoral diversity at high levels of disaggregation (e.g., 5-digits industrial classes), while negatively correlating with differences in specialization at low levels of disaggregation (e.g., 2-digits technological classes). Following this approach, "*related variety [would] be the indicator for Jacobs externalities because it measures the variety within each of two digit classes. It is expected that the economies arising from variety are especially strong between subsectors, as knowledge spills over primarily between firms selling related products*" (Frenken et al., 2009, p. 689).

The concept of cognitive proximity has been applied in regional analysis in order to capture the positive spillover effect stemming from sectoral diversity, which ultimately relates to the notion of cognitive proximity between actors. In this paper, a major step forward is made by extending the concept of cognitive proximity to the case of inter-regional knowledge flows. This leads to the definition of the concept of cross-regional cognitive proximity.

D1. Cross regional cognitive proximity	*Two regions are cognitively proximate if they enjoy a complementary set of skills and competences pertaining to a common knowledge base, characterizing a technological domain.*

In order to learn from each other, agents in distant regions must have compatible sets of skills (i.e. they have to share a common technological domain), within which sufficient cross-regional complementarity must exist. The extent to which a specific knowledge base matters within an

inter regional cognitive proximity depends on the relevance with respect to the overall knowledge present in the region.

In order to operationalize the concept of cross-regional cognitive proximity, we resort on regional patent data, which represent a good proxy for the knowledge profile of a region (see the Appendix for more details on the patent data).

Empirically, the common technological domain is approximated by a common specialization of pairs of regions into the same technological class (1 digit) of patents; potential for advancements is approximated by differentiation and complementarity in terms of specialization in sub-classes of patents (2 digits) in two regions. The higher the difference between the two regional shares of patents in 3-digit technological classes, the higher the complementarity between regions. The higher the share of 2-digit technological classes in the pair of regions analyzed, the higher the common knowledge base in the region. The latter should be adjusted for the difference in the size of the same technological domain in the two regions; the higher the difference in size, the lower the cognitive proximity. [4]

This definition can be easily translated into the following cross-regional cognitive proximity (between region *i* and *j*) indicator (cog_{ij}):

$$cog_{ij} = \sqrt{\sum_{s_{2d}=1}^{7}\left[\frac{\left(x_{s_{2d}i} * x_{s_{2d}j}\right)}{\left|x_{s_{2d}i} - x_{s_{2d}j}\right|}\left(\sum_{s_{3d}=1}^{m}\left(\left|x_{s_{3d}i} - x_{s_{3d}j}\right|\right)\right)\right]} \qquad (6)$$

This indicator is a positive function of the within 3-digits class variety (complementarity of knowledge among regions), a negative function of the two regions' overall patenting differences, and a positive function of the 2-digits specialization (common knowledge base between a pair of regions).

Finally, the whole indicator is square-rooted in order to reduce the variance of otherwise extreme data, typical of indicators based on patent counts as also suggested in the literature (Hollanders et al., 2009). As a final step, the values of *cog* have been normalized on the maximum, so that the indicator ranges from 0 to 1.

d) Social proximity

The social proximity matrix is based on the concept of social capital. Social capital has been used in order to formalize the cultural characteristics of a society facilitating interactions, reducing transaction costs, improving the ease with which knowledge spreads (Capello et al., 2011; Basile et al., 2012).

The idea behind this paper is that regions more similar in terms of their social capital exchange knowledge more easily. On the contrary, pairs of regions with different social capital face higher transaction costs in the process of absorbing, understanding and decoding external knowledge.

In this paper we follow Putnam's definition of social capital: "*Social capital here refers to features of social organization, such as trust, norms, and networks, that can improve the efficiency of society by facilitating coordinated actions* (Putnam et al. 1993, p. 167)". Given the paramount relevance of coordination in the complex research activities being explained in the present work, and the even higher relevance of coordination-enhancing societal characteristics, it follows that regions with different stocks of social characteristics may find it harder to structure complex R&D cooperation networks. At the same time, such impedance factor may ceteris

[4] This component of the indicator allows also to take into account possible extreme values of the regions' patenting specialization profiles, which happens as two regions have lo patenting intensities and have therefore very high shares of patenting in one specific 2-digits patent class, just because no other 2.digitsd class is covered.

paribus (viz. with similar absolute differences in social capital characteristics) be reduced as regions are simultaneously characterized by higher levels of social capital.

In order to construct the social distance matrix, first a measure of social capital must be built. The empirical definition of social capital is based on the seminal work by Putnam and coauthors (Putnam, 2000; Putnam et al., 1993). This measure is calculated by averaging out percentage scores to questions in the individual questionnaires administered in the 2000 wave of the European Values Study (henceforth, EVS).[5] For each of the theoretical domains of social capital originally identified as crucial elements of this concept in Putnam (2000), a proper proxy has been identified among EVS questions. Individual answers have next been aggregated at the European NUTS2 level.[6] In particular, the choice of the questions selected to represent the domains of the Putnam literature are shown in Table 1.

Table 1. Indicators of social capital used for the social proximity matrix

Domain	Question	Scale
Community organizational life	How often do you spend time in clubs and voluntary associations?	1 every week 2 once or twice a month 3 a few times a year 4 not at all
Engagement in public affairs	Do you participate in any form of social activity?	0-1
Community volunteerism	Do you take part to voluntary work in any community activity?	0-1
Informal sociability	Do you agree that “Most people can be trusted”	1 I trust them completely 2 I trust them a little 3 I neither trust nor distrust them 4 I do not trust them very much 5 I do not trust them at all

Once a proper indicator for each of Putnam’s social capital axes has been defined, and the economic rationale for the assessment of interregional differences in terms of social capital explained, the concept of cross-regional social proximity can be defined as follows:

D2. Social proximity — *Two regions are socially proximate if they enjoy similar sets of social values. In order to learn from each other, agents in different regions must have compatible sets of values.*

Empirically, social distance between regions *i* and *j* is computed in terms of normalized Euclidean distances between regional social capital characteristics with the formula:

$$soc_{ij} = \sqrt{\sum_{q=1}^{Q}\left(x_{qi} - x_{qj}\right)^2} \Big/ \sum_{q=1}^{Q}\left(x_{qi} + x_{qj}\right) \tag{7}$$

where x_q are the social capital indicators (with Q=4). The indicator ranges from 0 to 1.

The way in which the indicator is built guarantees that what matters in generating more/less social proximity between pairs of regions is the lower/higher distance in their stock of social values, whatever the stock of social capital is in the two regions. A pair of regions with a low stock of social capital but a similar social capital profile has a higher social proximity than a pair

[5] The EVS is a large scale longitudinal survey research project aiming at investigating fundamental value patterns among European citizens. Its main goal is to “*empirically uncover basic values, attitudes, and preferences of the European population and to explore the similarities, differences, and changes in these orientations*” (Halman 2001, p.1).

[6] More information on the validity of regional sampling in the EVS can be found in Caragliu (2010).

of regions with a higher stock of social capital but very different social capital profiles, which is exactly what we mean by social proximity.

e) Specialization proximity

In order to be inclined to cooperate for scientific purposes, regions must enjoy a compatible productive context. Similar specialization patterns in manufacturing sectors, in particular, are expected to increase the likelihood of research collaboration between pairs of regions. This leads to the following definition of cross-regional specialization proximity:

D3. Specialization proximity — *If two regions enjoy similar specialization in manufacturing industries, they are specialization-wise proximate. In order to learn from each other, regions must have compatible levels of specialization in manufacturing industries, where scientific activity is performed (i.e. they must have specialization similarities).*

In fact, it has been noted that "*Similarities in technological knowledge (…) facilitate technological learning as well as the anticipation of technological developments (…). Technological proximity between actors facilitates the acquisition and development of technological knowledge and technologies*" (Knoben and Overmans, 2006, p. 77).

It may be claimed that the concept of specialization proximity hides a certain similarity with the concept of cognitive proximity; industrial specialization can explain the technological performance of a region and therefore its knowledge profile. However, we believe specialization proximity adds something to the analysis, by:

- defining that research cooperation is easier among actors operating in similar industrial contexts;
- overcoming a limit of the proxy of knowledge used in the cognitive proximity measure. The use of patents captures formal knowledge but leaves aside all informal, tacit knowledge embedded in specific technological capabilities that can exist in a region.[7]

The matrix of cross-regional specialization proximity is operationalized as follows. In the first place, location quotients for manufacturing employment (Table 2 shows the NACE 2 classes used in this process) are calculated, with the EU27 as the reference value.

Table 2. Manufacturing industries used for the calculation of the specialization proximity matrix

Industry code	Industry name
DA	Food, beverages and tobacco
DBDC	Textiles and leather etc.
DFDGDH	Coke, refined petroleum, nuclear fuel and chemicals etc.
DL	Electrical and optical equipment
DM	Transport equipment
OM	Other manufacturing

Next, distances between pairs of regions with respect to the location quotients are calculated. Each entry of the specialization proximity matrix, therefore, takes on the following formula:

$$tech_{ij} = \frac{\sqrt{\sum_{q=1}^{Q}\left(LQ_{qi} - LQ_{qj}\right)^2}}{Q} \tag{8}$$

[7] This aspect would be completely overcome if data at industrially disaggregated level were available.

where LQ stands for region-specific Location Quotients, indices *i* and *j* indicate regions, while *Q* is the number of manufacturing industries here considered (in this case 6, as shown in Table 2), so that the indicator ranges from 0 to 1.

f) Control variables

The main determinants of scientific collaboration between regions used in this paper are of two types.

The first category of control variable captures the physical size and the size of the knowledge produced by the pair of regions, measured respectively as the average number of inhabitants of each pair of regions and the average R&D expenditure over GDP of each pair of regions. In order to overlap as perfectly as possible with the dependent variable, and reduce the potentially distorting impact of business cycles, the two size variables are calculated as averages of the yearly values in 1998-2002, exactly the period when FP5 projects were funded.

The second category of control variable measures the difference in the capability of the pair of regions to produce either knowledge or innovation, measured respectively as the absolute difference in R&D expenditure over GDP and in the share of firms performing product innovation according to the 2002-2004 wave of the Community Innovation Survey (CIS[8]) between the pairs of regions. These variables allow to control for differences that can explain the interest of regions for cooperation, irrespective of all other types of proximities to other regions. In fact, ceteris paribus, scientific cooperation is expected to take place among regions with similar research capabilities and with similar skills in translating knowledge into a commercialized and high value-added product. The inclusion of these two variables allows to further reduce the role played by geographical proximity in mediating scientific collaboration, by stressing the arguably negative contribution of large differences in R&D investment and product innovation intensity between pairs of regions on their likelihood to cooperate.

Measuring scientific collaboration with regions' co-participations in EU FP5 projects implies that a specific weight is given to trans-national cooperation, which is fostered by the very nature of Framework Programmes. In fact, the EU believes the FP should be best targeted towards selecting "*only objectives which are more efficiently pursued at the Community level by means of research activities conducted at that level*", while at the same time meeting the "*need to establish a "critical mass" in human and financial terms, in particular through the combination of the complementary expertise and resources available in the various Member States*".[9] As such, it is highly likely that scientific cooperation measured by FP5 projects is negatively associated to the fact that regions belong to the same country. We control for this issue with a dummy which takes on value 1 if indeed regions belong to the same country.

Table 3 shows finally the correlation between the four main proximity matrices. Not necessarily such concept of proximity are linked – in fact, in some cases they even correlate negatively. This first statistical analysis reinforces the case to add alternative proximity measures to the explanation of cross-regional research cooperation.

Table 3. Pearson's correlations between proximity measures

	Cognitive proximity	Specialization proximity	Social proximity	Geographical proximity
Cognitive	1			

[8] CIS data are not publicly available for European NUTS2 regions for all countries of the EU27. In this paper we use CIS vectors produced within the ESPON KIT project (http://www.espon.eu/main/Menu_Projects/Menu_AppliedResearch/kit.html). For details on the innovation data, see Capello et al. (2012).

[9] Sources of the cited sentences and further information on this point can be found at the FP5's information web site, at the URL http://cordis.europa.eu/fp5/src/criteria.htm (date of retrieval: Apr. 13, 2012).

proximity	(0.00)			
Specialization proximity	-0.0237	1		
	(0.45)	(0.00)		
Social proximity	-0.0122	0.0297	1	
	(0.00)	(0.00)	(0.00)	
Geographical proximity	-0.0164	-0.1810	-0.0947	1
	(0.00)	(0.00)	(0.00)	(0.00)

Note: Standard errors in parentheses.

5. Empirical results

5.1 The joint effect of geographical and non-geographical proximities on scientific collaboration

In this Section the results of estimating eq. (4) with the use of the indicators described in Section 4 are presented. Each Sub-Section aims at answering one of the research questions described in Section 2.

Across all estimates, in this Section robust standard errors are used, in order to correct for potential heteroskedasticity in the data.

Our baseline estimates (Table 4) start from eq (4); first, in column 1 and 2, we include only two measures of the sheer size of the regions potentially involved in scientific collaboration (namely, the average percentage expenditure in R&D over GDP in both regions belonging to each regions' pair, and their average populations), along with the dummy variable indicating when regions belong to the same Country; these variables are maintained throughout the estimates.

Next (columns 2-7), we progressively include all distance measures described in the previous Section (in the order geographical distance-cognitive proximity-specialization proximity -social proximity -R&D difference-product innovation difference). Finally, in columns 8 and 9 we also verify whether our results are driven by the use of linear log transformations or instead hold with the use of count models.

The results show remarkable significance, with all signs in line with theoretical expectations, and stability across all specifications. The only (expected) exception is the geographical distance parameter, which decreases with the inclusion of more measures of proximity (Figure 2). This is in line with the theoretical rationale already pointed out in Ertur and Koch (2011): in this literature, geographical space is at best a good proxy for the real underlying mechanisms driving the diffusion of knowledge. While its importance cannot be ignored, the magnitude of its true impact should probably be revised with the inclusion of richer ways to conceive of space.

The results are not driven by the use of standard OLS techniques. Columns 8 and 9 show in fact that negative binomial estimates obtain qualitatively comparable results, with similar signs (although with different magnitudes, which is in line with analogous studies).[10] Column 9, besides, takes also into account the existence of zero-observations in the data set, i.e. regions' pairs which never collaborate during the 1998-2002 FP cycle. In this case, the zero-inflation problem is modeled with a zero-inflated negative binomial technique, which fits two models, a first-step regression which explains the probability that two regions embark in scientific cooperation; and next, a second (viz. main) equation, which identifies the partial correlations between actual scientific cooperation (i.e. nonzero observations in the dependent variable) and right-hand side determinants.

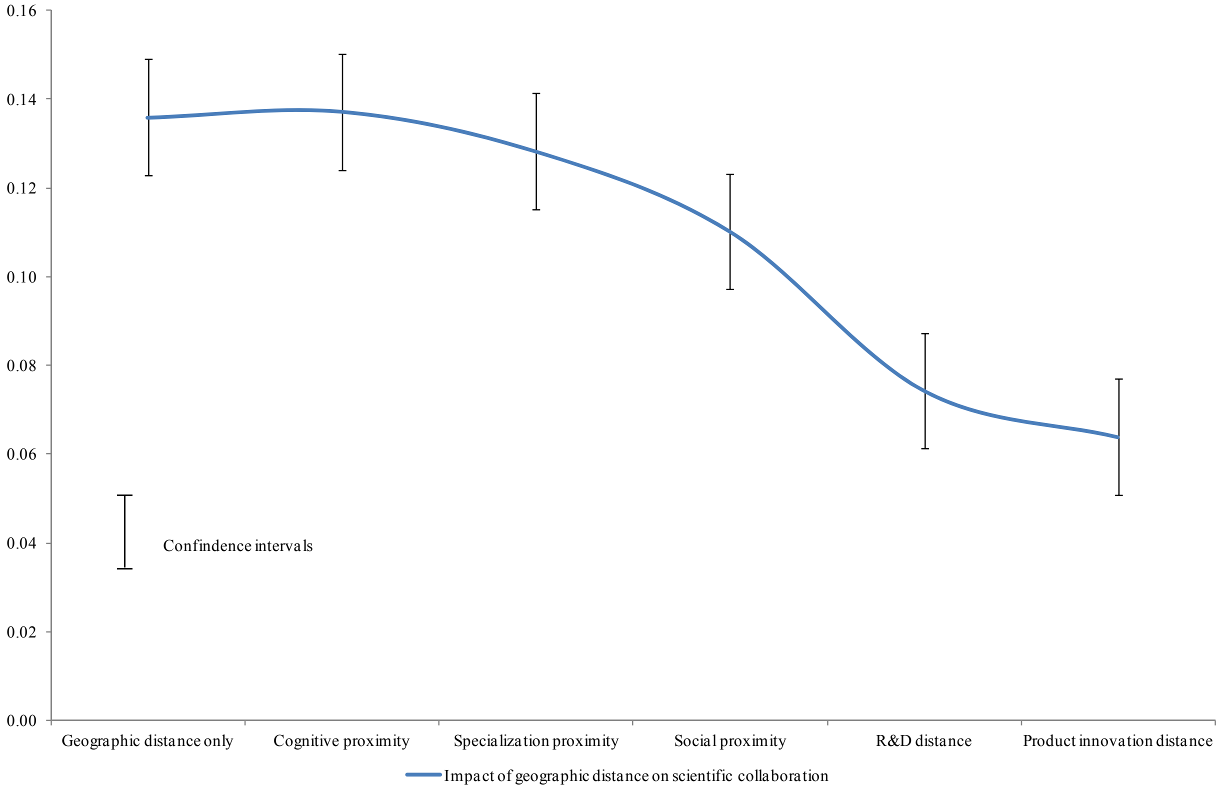
[10] Similar findings have been obtained with the use of Poisson and Zero-Inflated Poisson regressions; results are available upon request.

Table 4. Empirical results – Linear estimates

Dep. variable	*(Log of) scientific collaboration*								
Estimation method	OLS	OLS	OLS	OLS	OLS	OLS	OLS	Negative Binomial	Zero-Inflated Negative Binomial
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	-0.07	0.96***	1.05***	1.12***	1.16***	0.62***	0.50***	-4.44***	-2.78***
	(0.08)	(0.11)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.23)	(0.21)
Cooperation determinants									
Size of knowledge produced (average R&D expenditure/GDP in region pair)	0.92***	0.89***	0.91***	0.91***	0.91***	0.97***	0.96***	0.95***	0.57***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Physical size of region (average population levels in region pair)	0.02	0.03**	0.05***	0.05***	0.05***	0.04***	0.05***	0.40	2.18***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.75)	(0.67)
Proximity measures									
Geographical distance	-	-0.14***	-0.14***	-0.13***	-0.11***	-0.07***	-0.06***	-0.0000454**	-0.0000293***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Cognitive proximity	-	-	0.04***	0.04***	0.04***	0.03***	0.03***	0.06***	0.02***
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
Specialization proximity	-	-	-	1.03***	1.02***	0.97***	0.93***	2.40***	2.00***
				(0.01)	(0.11)	(0.11)	(0.11)	(0.21)	(0.19)
Social proximity	-	-	-	-	0.96***	0.93***	0.88***	2.00***	1.43***
					(0.08)	(0.08)	(0.08)	(0.15)	(0.14)
Difference in knowledge production between pairs of regions	-	-	-	-	-	-0.19***	-0.19***	-1.86***	-1.00***
						(0.01)	(0.01)	(0.05)	(0.05)
Difference in innovation production between pairs of regions	-	-	-	-	-	-	-0.05***	-0.24***	-0.29***
							(0.01)	(0.05)	(0.05)
Dummy, equals 1 if regions belong to the same Country	0.19***	0.01	-0.003	-0.02	-0.09***	-0.07**	-0.08**	-0.12***	-0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Robust standard errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs. (nonzero)	21,012	21,011	20,646	20,646	20,646	20,479	20,332	34,717	34,717 (21,012)
R squared	0.18	0.18	0.19	0.19	0.19	0.22	0.22	0.21	0.21
Joint F test (OLS)/Wald Chi Square test (Poisson)	1533.01***	1175.51***	951.77***	816.98***	728.00***	753.81***	662.12***	7570.55***	3542.69***

*Note: Standard errors in parentheses. * = significant at the 90% level; ** = significant at the 95% level; *** = significant at the 99% level.*

Figure 2. Values of the estimated parameter for the geographical distance measure, as additional proximity measures are included.



Since our results hold with the use of count techniques, and given the following two research questions, in the remain of the paper we opt for classical OLS estimates.

5.2 Synergies between geographical and non-geographical proximities

In this Subsection, we deal with the second research question (Table 5). In particular, we verify whether synergies exist between different forms of proximity: do regions cooperate more easily with regions being not only spatially, but also cognitively, specialization-wise, and socially similar? This question is answered by interacting the main measures of proximity of interest (namely, cognitive, specialization, and social proximity) with the measure of geographical distance.

Table 5 shows the results of estimating the main model with the inclusion of each of the interacted terms: column 1 shows the synergic interaction between geographical distance and cognitive proximity, column 2 between geographical distance and specialization proximity, column 3 between geographical distance and social proximity.

The first column shows that all three components of the interaction terms display high significance. However, as usual with interaction terms between continuous variables, the results for individual parameter estimates are interesting for different values of the interaction term. For this reason we resort on graphical analysis.

Figures 3, 4, and 5 show the marginal effects of, respectively, cognitive, specialization, and social proximity as geographical distance increases.

Figure 3 shows an interesting and counter-intuitive result. For low levels of geographical distance, the effect of cognitive proximity on scientific collaboration is negative, while when regions are located far away, the negative effect of cognitive proximity on cooperation decreases, and even become positive, witnessing that cognitive proximity is required for scientific cooperation when a strong geographical distance is in place. This is in line with similar findings obtained on the basis of a large micro data base in Boschma et al. (2009).

Table 5. Empirical results – Interaction estimates

Dep. variable	*Log of scientific collaboration*		
Model	(1)	(2)	(3)
Constant	-0.56** (0.28)	0.15 (0.15)	0.20 (0.33)
Cooperation determinants			
Size of knowledge produced (average R&D expenditure/GDP in region pair)	0.96*** (0.01)	0.96*** (0.01)	0.96*** (0.01)
Physical size of region (average population levels in region pair)	0.05*** (0.01)	0.05*** (0.01)	0.05*** (0.01)
Proximity measures			
Geographical distance	0.11*** (0.04)	0.01 (0.01)	-0.13*** (0.03)
Cognitive proximity	-0.57*** (0.13)	0.07*** (0.01)	0.06*** (0.01)
Specialization proximity	0.93*** (0.11)	-3.56*** (0.84)	0.93*** (0.11)
Social proximity	0.88*** (0.08)	0.89*** (0.08)	3.35*** (0.79)
Difference in knowledge production between pairs of regions	-0.19*** (0.01)	-0.19*** (0.01)	-0.19*** (0.01)
Difference in Innovation production between pairs of regions	-0.05*** (0.01)	-0.04*** (0.01)	-0.05*** (0.01)
Interaction geographical distance-cognitive proximity	0.09*** (0.02)	-	-
Interaction geographical distance-specialization proximity	-	0.67*** (0.13)	-
Interaction geographical distance-social proximity	-	-	-0.35*** (0.11)
Dummy, equals 1 if regions belong to the same Country	-0.09*** (0.02)	-0.07*** (0.03)	-0.15*** (0.03)
Robust standard errors	Yes	Yes	Yes
Number of obs.	20,332	20,332	20,332
R squared	0.22	0.22	0.22
Joint F test	601.91***	603.01***	595.72***

*Note: Standard errors in parentheses. * = significant at the 90% level; ** = significant at the 95% level; *** = significant at the 99% level*

Figure 3. Marginal effects of cognitive proximity on scientific collaboration for different geographical distances

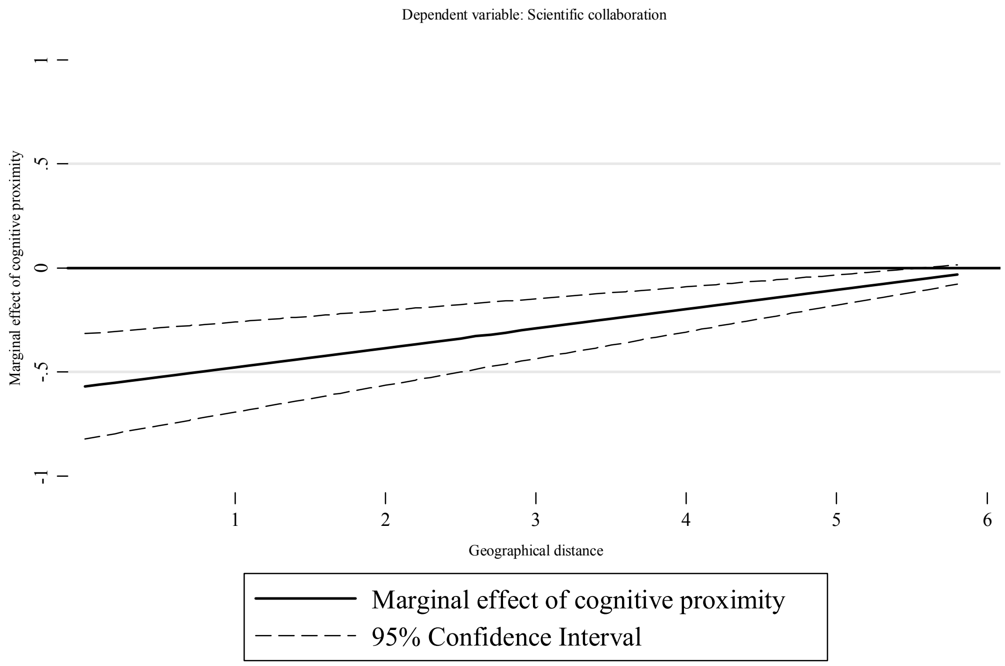
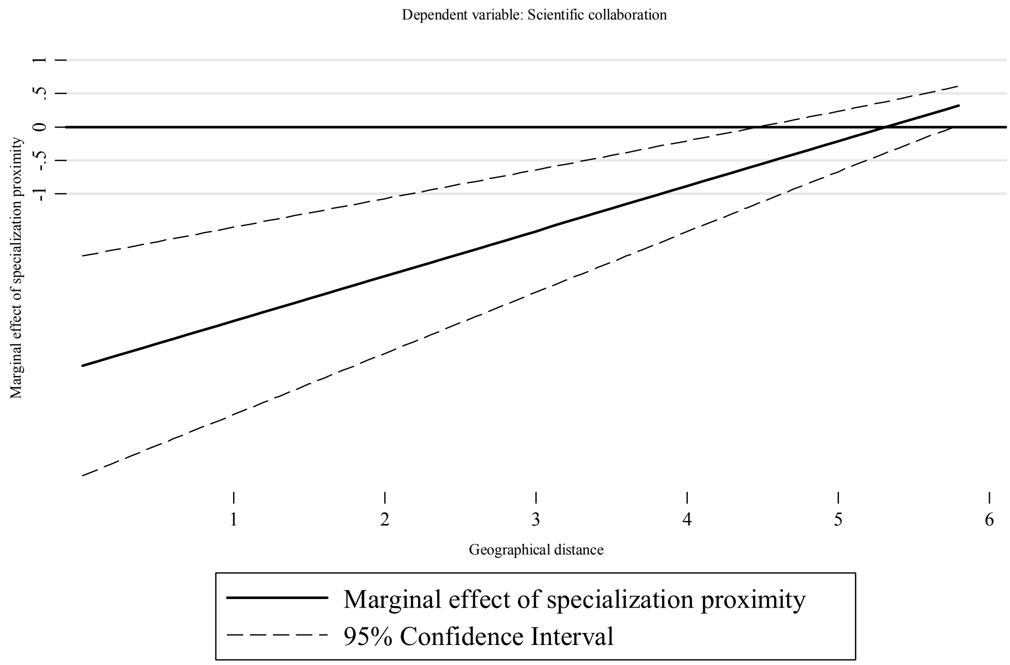


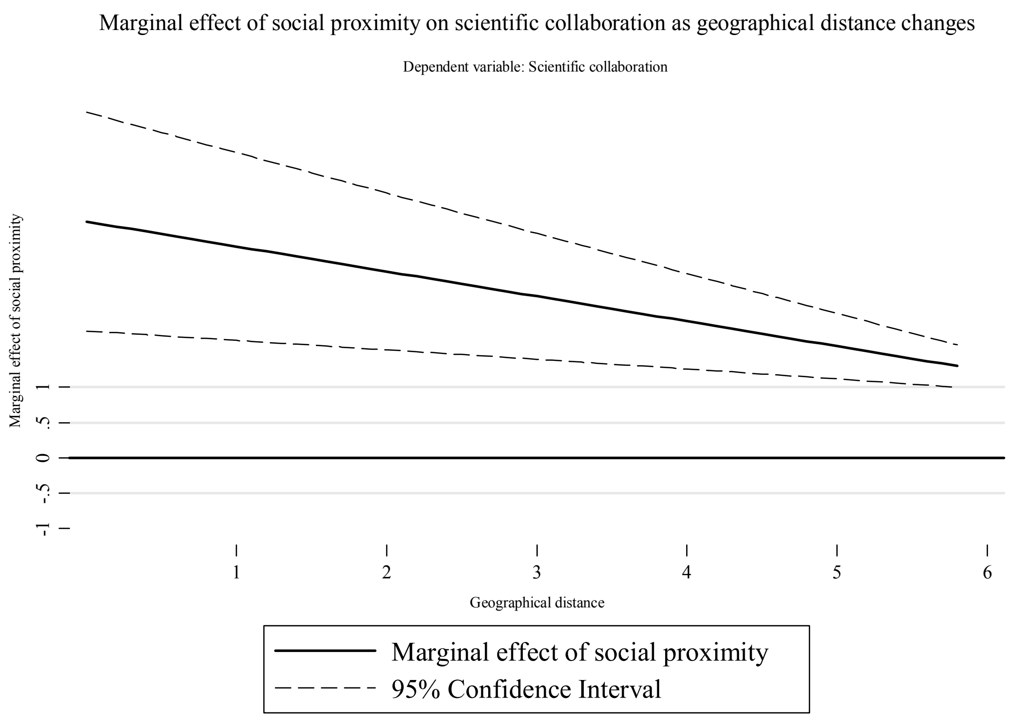
Figure 4 shows, for given levels of specialization distance, its marginal effect on scientific collaboration as geographical distance increases. This Figure can be interpreted along the same lines of Figure 3; the effect of specialization proximity on scientific cooperation is positive only when regions are located far away. Similarities in industrial profiles, and therefore in technological knowledge, are required for scientific cooperation in the case of distant locations. These similar results are welcome given the similar conceptual interpretation of the cognitive and specialization proximities.

Figure 4. Marginal effects of specialization proximity on scientific collaboration for different geographical distances



Finally, Figure 5 displays the impact of social proximity on scientific collaboration as geographical distance increases. In this case the impact of social proximity on scientific collaboration remains positive throughout the range where geographical distance is defined; this positive impact, nevertheless, decreases in magnitude as geographical distance between regions engaging in scientific cooperation increases. This implies that, even at large geographical distances, regions being characterized by different social values find it difficult to collaborate for scientific purposes.

Figure 5. Marginal effect of social proximity on scientific collaboration for different geographical distances



5.3 Nonlinearities in the impact of proximity on scientific collaboration

This final Subsection verifies whether nonlinear effects exist in the role of proximities on scientific collaboration. In order to maintain throughout the paper similar techniques, possible nonlinearities are inspected by means of robust OLS estimates of not only the linear, but also the squared parameters. Results of the third empirical exercise are shown in Table 6.

The results are organized as follows. The baseline model in Section 5.1, Table 4, Column 8 is for easing the cross-readability of our results reproduced in Column 1 in Table 6. In the rest of the Table we add the squares of all the main distance/proximity measures, one per column. For instance, Column 2 adds the square of geographical distance, column 3 the square of cognitive proximity, and so forth.

Results present interesting patterns:

- for geographical distance, instead, the squared parameter estimate displays a positive value, which implies a full-fledged distance decay effect, with a negative, but less than proportional, impact of geographical distance on scientific collaboration. Decreasing returns seem to be present;
- for cognitive and social proximity, the squared parameter estimates display negative and significant values, which, along with the positive estimate for the linear parameter, implies that scientific collaboration increases, but at a decreasing rate, with all abovementioned proximity measures, displaying therefore decreasing returns;

- finally, the specialization proximity squared parameter turns out to be positive and extremely large; this implies that the probability that two regions cooperate for scientific reasons increases, and more than proportionally so, as their specialization profiles get more similar.

Table 6. Empirical results – Nonlinear estimates

Dep. variable	*Log of R&D collaboration*				
Model	(1)	(2)	(3)	(4)	(5)
Constant	0.50***	4.03***	3.84***	3.88***	3.89***
	(0.12)	(0.40)	(0.40)	(0.39)	(0.39)
R&D determinants					
R&D intensity (average R&D expenditure/GDP in region pair)	0.96***	0.97***	0.97***	0.97***	0.97***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Size of region (average population levels in region pair)	0.05***	0.05***	0.05***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Proximity measures					
Size of knowledge produced (average R&D expenditure/GDP in region pair)	-0.06***	-1.10***	-1.11***	-1.05***	-1.04***
	(0.01)	(0.12)	(0.12)	(0.11)	(0.11)
Physical size of region (average population levels in region pair)	-	0.08***	0.08***	0.08***	0.08***
		(0.01)	(0.01)	(0.01)	(0.01)
Cognitive proximity	0.03***	0.06***	0.21***	0.21***	0.21***
	(0.01)	(0.01)	(0.05)	(0.05)	(0.05)
Square of cognitive proximity	-	-	-0.08	-0.08	-0.08
			(0.02)	(0.02)	(0.02)
Specialization proximity	0.93***	0.93***	0.94***	3.95***	3.94***
	(0.11)	(0.11)	(0.11)	(0.24)	(0.24)
Square of specialization proximity	-	-	-	6.88***	6.86***
				(0.47)	(0.47)
Social proximity	0.88***	0.90***	0.92***	0.88***	1.25***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.77)
Square of social proximity	-	-	-	-	-0.16***
					(0.03)
Difference in knowledge production between pairs of regions	-0.19***	-0.19***	-0.19***	-0.18***	-0.18***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Difference in innovation production between pairs of regions	-0.05***	-0.05***	-0.05***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Dummy, equals 1 if regions belong to the same Country	-0.08**	-0.16***	-0.16***	-0.20***	-0.21***
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Robust standard errors	Yes	Yes	Yes	Yes	Yes
Number of obs.	20,332	20,332	20,332	20,332	20,332
R squared	0.22	0.22	0.23	0.23	0.23
Joint F test	662.12***	603.82***	550.99***	526.62***	486.32***

*Note: Standard errors in parentheses. * = significant at the 90% level; ** = significant at the 95% level; *** = significant at the 99% level.*

Such findings can be summarized in a qualitative way as shown in Figures 6-9, where on the x-axes the measures of geographical distance, and cognitive, specialization, and social proximity are plotted, while on the y-axes we represent the marginal effects of, respectively, geographical distance, and cognitive, specialization, and social proximity on scientific collaboration.

Figure 6. Marginal effect of geographical distance on scientific collaboration as geographical distance increases

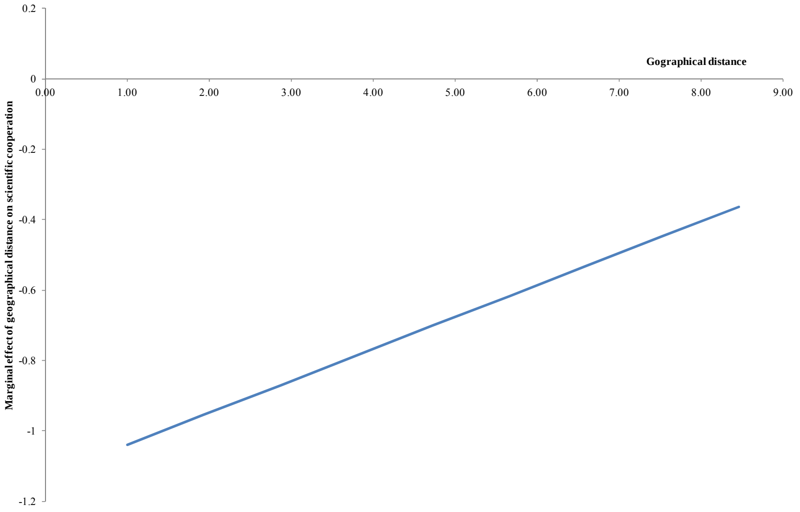


Figure 7. Marginal effect of cognitive proximity on scientific collaboration as cognitive proximity increases

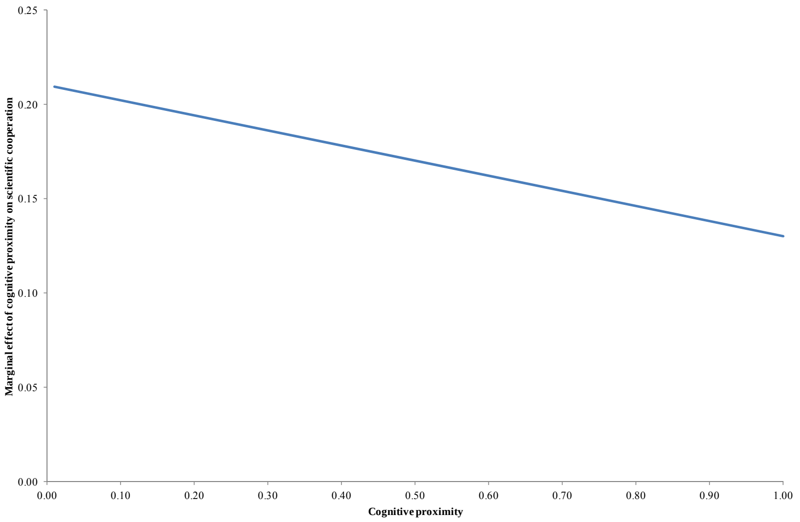


Figure 8. Marginal effect of specialization proximity on scientific collaboration as specialization proximity increases

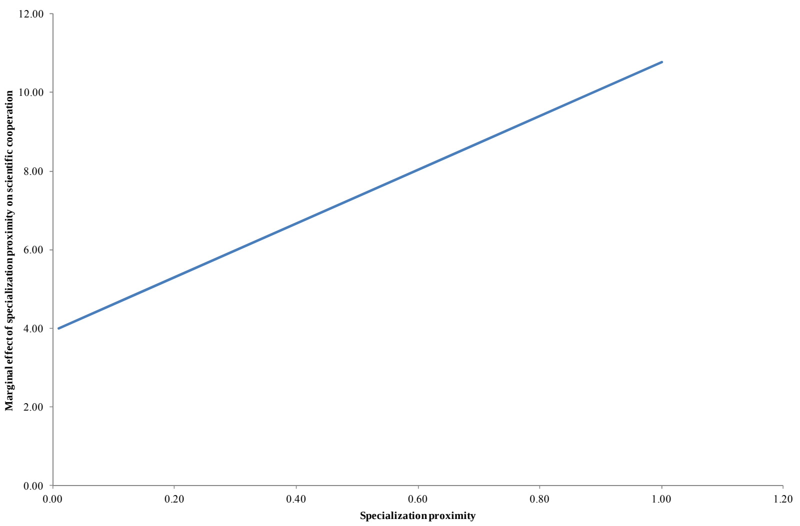
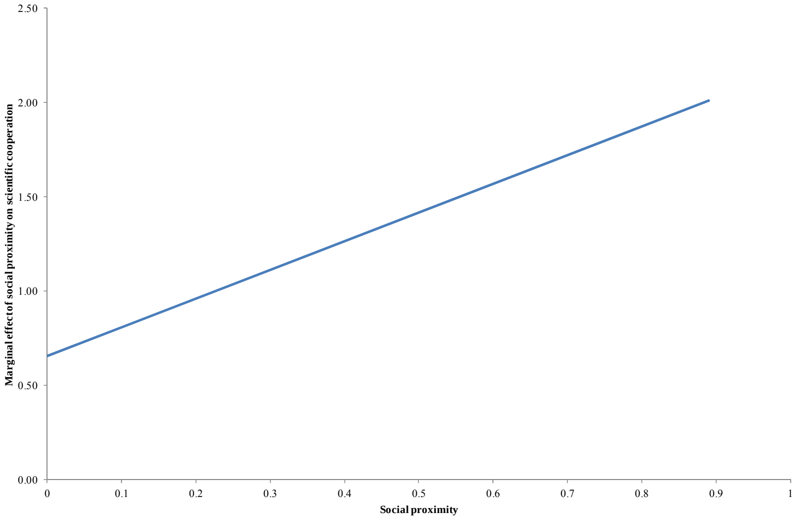


Figure 9. Marginal effect on scientific collaboration of social proximity as social proximity increases



6. Conclusions

This paper enters the debate on the ways space can be conceived of in modeling economic interactions, and provides empirical evidence in three main directions.

Firstly, evidence is provided on the role of a large number of distance concepts, which widens the evidence on the distance-decay effects affecting scientific cooperation, in line with previous research on this topic.

Secondly, the results point towards the existence and relevance of synergic effects between different forms of space. In particular, whilst social proximity impacts positively scientific cooperation, with such impact decreasing in magnitude as geographical distance increases, the results on cognitive and specialization proximity suggest that some form of complementarity seems to exist with geographical distance. At low geographical distance, some degree of cognitive and specialization distance seem to be bearable for the purpose of scientific cooperation, with regions filling the cognitive and specialization gap exactly by means of such low geographical distance. When, however, geographical distance increases, such bridge ceases to exist and work, which implies that in order to cooperate, regions must be cognitively and specialization-wise close.

Thirdly, and lastly, our results point, for the first time to our knowledge, towards the existence of non-linearities in the impacts of non-geographical distance on scientific cooperation. While such results are *per se* interesting, they call for further research, possible with the use of partially, or fully, nonlinear estimators, thereby allowing the mapping of such nonlinear effects.

Far from being merely academic, these results call for profound policy implications. Currently, the future EU research Agenda is being shaped, and discussed at the highest policy levels. Ignoring possible synergies between distance decay effects, for instance, may imply that the (formal and informal) rules for shaping transnational research groups may be mislead, and ultimately ineffective. Along the same lines, ignoring nonlinearities in distance-decay effects may lead to over-or under-funding of specific research lines, or even wrongly assume that multidisciplinarity can afford the formation of large transnational research groups with various scientific backgrounds, ultimately leading to a cost-ineffective provision of research funds.

This last point, in connection with the decrease in research funds in most countries, because of the ongoing financial crisis, may lead to the failure to reach the “more with less” goal that most policymakers currently seem to have as a major mindset.

References

Agrawal, A., Kapur, D., and McHale, J. (2008). "How do spatial and social proximity influence knowledge flows? Evidence from patent data", *Journal of Urban Economics*, 64 (2): 258-269.

Akerlof, G. (1997). "Social distance and social decisions", *Econometrica*, 65 (5): 1005-1027.

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

Anselin, L., Varga, A. and Acs, Z. (1997). "Local geographic spillovers between university research and high technology innovations", *Journal of Urban Economics,* 42 (3): 422-448.

Antonelli, C. (2003). "*The economics of innovation, new technologies and structural change*", Studies in global competition 14, London-New York: Routledge.

Autant-Bernard, C., and LeSage, J. (2011). "Quantifying knowledge spillovers using spatial econometric models", *Journal of Regional Science,* 51 (3): 471-496.

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 (3): 495-519.

Autant-Bernard, C. (2001). "The geography of knowledge spillovers and technological proximity", *Economics of Innovation and New Technology*, 10 (4): 237-254.

Basile, R., Capello, R., and Caragliu, A. (2012). "Technological interdependence and regional growth in Europe: proximity and synergy in knowledge spillovers", forthcoming in *Papers in Regional Science*.

Belot, M., and S. Ederveen (forthcoming). "Cultural and institutional barriers in migration between OECD countries", *Journal of Population Economics*, available online under DOI 10.1007/s00148-011-0356-x.

Boschma, R.A. (2005). "Proximity and innovation: a critical assessment", *Regional Studies*, 39 (1): 61-74.

Boschma, R. A., Eriksson, R., and Lindgren, U. (2009). "How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity", *Journal of Economic Geography*, 9 (2): 169-190.

Boschma, R. A., and Iammarino, S. (2009). "Related variety, trade linkages, and regional growth in Italy", *Economic Geography*, 85 (3): 289-311.

Breschi, S. and Lissoni, F. (2001). "Localised knowledge spillovers vs. innovative milieux: knowledge "tacitness" reconsidered", *Papers in Regional Science*, 80 (3): 255-273.

Capello, R. (2009). "Spatial spillovers and regional growth: a cognitive approach", *European Planning Studies,* 17 (5): 639-658.

Capello, R., Caragliu, A., and Lenzi, C. (2012). "Is innovation in cities a matter of knowledge-intensive services? An empirical investigation", *Innovation- The European Journal of Social Science Research*, 25 (2): 151-174.

Capello, R., Caragliu, A., and Nijkamp, P. (2011). "Knowledge and social capital: increasing returns in regional growth", *Studies in Regional Science*, 41 (1): 95-116.

Caragliu, A. (2010). "*Innovation, territorial capital and regional growth*", Ph.D. thesis, Politecnico di Milano.

Cowan, R., Jonard, N., and Zimmermann, J. B. (2007). "Bilateral collaboration and the emergence of innovation networks, *Management Science,* 53 (7):1051-1067.

Coe, D. T., and Helpman. E. (1995). "International R&D spillovers", *European Economic Review*, 39 (5): 859-887.

Ertur, C, and Koch, W. (2011). "A contribution to the Schumpeterian growth theory and empirics", *Journal of Economic Growth*, 16 (3): 215-255.

Ertur, C. and Koch, W. (2007). "Growth, technological interdependence and spatial externalities: theory and evidence", *Journal of Applied Econometrics*, 22 (6): 1033-1062.

Frenken, K., Oort, F. v., and Verburg, T. (2009). "Related variety, unrelated variety and regional economic growth", *Regional Studies,* 41 (5): 685-697.

Frenken, K., Ponds, R., and Oort, F. van (2010). "The citation impact of research collaboration in science-based industries: a spatial-institutional analysis", *Papers in Regional Science*, 89 (2): 351-371.

Groot, H. L. F. de, Linders, G. J., Rievtveld, P. and Subramanian, U. (2004). "The institutional determinants of bilateral trade patterns", *Kyklos*, 57 (1): 103-124.

Halman, L. (2001). "*The European Values Study: a third wave. Source book of the 1999/2000 European Values Study surveys*", Tilburg University: EVS-WORC.

Hollanders, H., Tarantola, S., and Loschky, A. (2009). "*Regional Innovation Scoreboard - Methodology Report*", retrieved online on June 19, 2012 at the URL http://www.proinno-europe.eu/page/regional-innovation-scoreboard.

Jaffe, A. (1989). "Real effects of academic research", *American Economic Review* 79 (5): 957-970.

Jaffe, A. (1986). "Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value", *American Economic Review*,76 (5): 984-1001.

Jaffe, A., Trajtenberg, M., and Henderson, R. (1993). „Geographic localization of knowledge spillovers as evidenced by patent citations", *The Quarterly Journal of Economics*, 108 (3): 577-598.

Knoben, J., and Overmans, L.A.G. (2006). "Proximity and inter-organizational collaboration: a literature review", *International Journal of Management Reviews*, 8 (2): 71-89.

Lankhuizen, M. de Groot, H. L. F., and Linders, G. J. M. (2011). "The trade-off between foreign direct investments and exports: the role of multiple dimensions of distance", *The World Economy*, 34 (8): 1395-1416.

MacGarvie, M. (2005). "The determinants of international knowledge diffusion as measured by patent citations", *Economics Letters,* 87 (1): 121-126.

Maggioni, M. and Uberti, E. (2009). "Knowledge networks across Europe: which distance matters?", *Annals of Regional Science*, 43 (3), 691-720.

Maggioni M., Nosvelli, M., and Uberti, E. (2007). "Space versus networks in the geography of innovation: a European analysis". *Papers in Regional Science* 86 (3): 471-493.

Mora, T., and Moreno, R. (2010). "Specialisation changes in European regions: the role played by externalities across regions", *Journal of Geographical Systems*, 12 (3): 311-334.

Nooteboom, B. (2000). "*Learning and innovation in organizations and economies*", Oxford: Oxford University Press.

Padilla-Pérez, R. (2008). "A regional approach to study technology transfer through foreign direct investment: the electronics industry in two Mexican regions", *Research Policy,* 37 (5): 849-860.

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

Putnam, R. D. (2000). "Bowling alone", New York (NY): Simon and Schuster.

Putnam RD, Robert L, and Nanetti R. (1993). "*Making democracy work: civic traditions in modern Italy*", Princeton (NJ): Princeton University Press.

Putnam RD, Robert L, and Nanetti R. (1993). "*Making democracy work: civic traditions in modern Italy*", Princeton (NJ): Princeton University Press.

Scherngell, T. and Barber, M. J. (2011). "Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU Framework Programme", *The Annals of Regional Science*, 46 (2): 247-266.

Scherngell, T. and Barber, M. J. (2009). "Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU Framework Programme", *Papers in Regional Science*, 88 (3): 531-546.

Spolaore, E. and Wacziarg, R. (2009). "The diffusion of development", *The Quarterly Journal of Economics,* 124 (2): 469-529.

Thompson, P. (2006). "Patent citations and the geography of knowledge spillovers: evidence from inventor- and examiner-added citations", *The Review of Economics and Statistics*, 88(2): 383-388.

Tinbergen, J. (1962). "*Shaping the world economy; suggestions for an international economic policy*", Appendix VI, New York (NY): The Twentieth Century Fund.

Uzawa, H. (1962). "Production functions with constant elasticities of substitution", *Review of Economic Studies*, 29 (4): 291-299.

van Pottelsberghe de la Potterie, B., and Lichtenberg, F. (2001). "Does Foreign Direct Investment transfer technology across borders?", *Review of Economics and Statistics*, 83 (3): 490-497.

Appendix

The 2-digits and 3-digits classes have the following summary statistics for our sample of 264 regions (Table A1 and Table A2):

Table A1. Summary statistics for the regional specialisation in 2digits classes.

Variable	Obs	Mean	Std. Dev.	Min	Max
Share patents class 1	264	17%	13%	0%	73%
Share patents class 2	264	15%	11%	0%	90%
Share patents class 3	264	12%	10%	0%	89%
Share patents class 4	264	11%	11%	0%	87%
Share patents class 5	264	15%	10%	0%	69%
Share patents class 6	264	19%	14%	0%	100%
Share patents class 7	264	12%	11%	0%	100%

Source: Authors' calculations.

Table A2. Summary statistics for the regional specialisation in 3digits classes.

2digits class	Variable	Obs	Mean	Std. Dev.	Min	Max
1	share3d_1	264	31.16%	25.09%	0%	100%
1	share3d_2	264	9.44%	10.66%	0%	100%
1	share3d_3	264	24.52%	21.60%	0%	100%
1	share3d_4	264	19.19%	21.54%	0%	100%
1	share3d_5	264	6.18%	12.68%	0%	100%
2	share3d_6	264	12.27%	17.35%	0%	100%
2	share3d_7	264	42.04%	25.54%	0%	100%
2	share3d_8	264	34.18%	25.91%	0%	100%
2	share3d_9	264	2.37%	8.19%	0%	100%
3	share3d_10	264	19.66%	20.61%	0%	100%
3	share3d_11	264	20.02%	20.80%	0%	100%
3	share3d_12	264	19.40%	18.90%	0%	100%
3	share3d_13	264	12.89%	15.00%	0%	100%
3	share3d_14	264	18.97%	19.68%	0%	100%
4	share3d_15	264	27.10%	22.03%	0%	100%
4	share3d_16	264	46.82%	27.62%	0%	100%
4	share3d_17	264	14.67%	18.26%	0%	100%
5	share3d_18	264	19.27%	16.89%	0%	100%
5	share3d_19	264	28.35%	21.05%	0%	100%

5	share3d_20	264	12.00%	15.42%	0%	100%
5	share3d_21	264	22.36%	18.81%	0%	100%
5	share3d_22	264	7.80%	13.19%	0%	100%
6	share3d_23 |	264	13.58%	12.65%	0%	63%
6	share3d_24	264	14.28%	17.55%	0%	100%
6	share3d_25	264	10.16%	13.42%	0%	100%
6	share3d_26	264	19.80%	17.27%	0%	100%
6	share3d_27	264	30.44%	22.95%	0%	100%
6	share3d_28	264	3.02%	10.09%	0%	100%
7	share3d_29 |	264	44.86%	28.01%	0%	100%
7	share3d_30	264	41.88%	27.29%	0%	100%

Source: Authors' calculations.