World Journal of Management
Vol. 7. No. 1. March 2016 Issue. Pp. 107 – 128

New Insights on the Relationships between Geographic and Institutional Distance in Research Collaborations: A Long Period Analysis

Rosamaria d'Amore[1] and Roberto Iorio[2]

This paper analyses the relationship between institutional distance and geographic distance in scientific collaborations, evaluating the possible changes when a long period (sixteen years: from 1990 to 2005) is taken into consideration. It also discusses the use of some alternative measurements of institutional distance. The main result, obtained by analysing the publications of the Italian biotech firms, is that international publications present a higher institutional distance than national papers, particularly in the early years, while there is no significant difference in institutional distance between regional and extra-regional papers, suggesting that opposite incentives are in action at different geographic scales and in different periods.

Keywords: university-industry collaboration; co-autorships; knowledge flows; scientific and technological policy.

JEL classifications: L14; O31; O32; R12

1. Introduction

It is widely acknowledged that in the modern knowledge-based economy, particularly in the science-based sectors, innovation usually derives from the collaboration of different agents, often located in different regions or even countries and often belonging to different institutions (OECD, 1996). The patterns of collaboration across regions or countries, the ease afforded by modern information technology for distant collaborations and the persistent importance of vis-à-vis collaboration are themes widely investigated in literature (Katz, 1994; Liang and Zhu, 2002; McKelvey, Alm and Riccaboni, 2003). On the other hand, an abundant stream of literature exists on the difficulties deriving from collaboration among different institutions, particularly between university and industry, which have different research goals and incentive structure (Dasgupta and David, 1994; Frenken and Van Oort, 2004). In any case, collaboration among different innovative institutions happens in several different ways (D'Este and Patel, 2007) and it may be fruitful in terms of quality because of the existence of complementarities (Bonaccorsi and Thoma, 2007; Iorio, Labory and Paci, 2012). According to Noteboom et al. (2009), cognitive distance provides the basis for resource heterogeneity across firms; therefore, it has a positive impact on innovation, at least up to a certain point.

[1]University of Salerno, Department of Economics and Statistics, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy, rmdamore@unisa.it
[2] University of Salerno, Department of Economics and Statistics, Via Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy, riorio@unisa.it

Relatively less explored is the theme of the relationship between these two kinds of "distance", the physical and institutional one, in research collaborations. Analysing this relationship is the main goal of this paper. This issue is, therefore, relevant, as a more articulated view about it may help to better understand how knowledge flows among innovative agents and to adopt better and more selective policy measurements.

The phenomenon of research collaboration is observed through the lens of co-authorship of scientific publications in the Italian biotech sector. This sector is particularly suitable for a study about research collaborations involving different institutions because it relies mostly on inter-organisational collaborations. As Powell, Koput and Smith-Doerr (1996) argue, in the biotech sector the locus of innovation will be found in networks rather than in individual firms. There are many organizations where it is possible to find the knowledge and expertise useful for the firm: the universities, the research centres, the hospitals. The new knowledge generated by these collaborations not only takes the form of industrial innovations, but it is often disclosed through the scientific publications: research collaborations often generate co-authored publications. Over two-thirds of even formal alliance partners in this field also appear as partners in scientific publications (Gittelman, 2005) and there is a close link between successful patents and scientific publications (Gittelman and Kogut, 2003; Murray and Stern, 2007).

The analysed database includes the publications done by the Italian biotech firms from 1990 to 2005. The institutions to which the authors of the publications belong are classified into four categories: firms, universities, hospitals and research centres. Their localisation is registered too, such that it is possible to identify the geographical extension of the collaborations, distinguishing among regional, extra-regional and international publications. As this research deals with institutional distance, it is important to measure it according to the existing literature (see, for instance, Ponds, Van Oort and Frenken, 2007). A collaboration between two institutions of the same kind (e.g. two firms or two universities) is characterised by the absence of institutional distance; a collaboration between two institutions of different kinds (e.g. one firm and one university; one hospital and one research centre) implies the presence of institutional distance. This research tries to go beyond this simple dichotomous distinction, suggesting some measurements based on the assumption that an increase in the variety of institutions involved in a collaboration increases the institutional distance. Once we establish the measurement of institutional distance, we will study its relationship with the geographic distance, verifying if and how the institutional distance varies when papers belonging to different geographic scales are compared. Thanks to a multivariate analysis, it is also possible to control for other factors that may have an effect on the institutional distance and geographic distance and on their relationship—the nature of the research, if basic or applied, the time trends and possible systematic differences in the publication practices of the firms. The long considered period also lets us observe the evolution overtime of institutional distance and geographic distance, and of their relationship and of other significant effects.

The same database used for this analysis has been analysed in previous studies (d'Amore et al., 2013; d'Amore, Iorio and Stawinoga, 2016), which have the same basic research question: what kind of relationship exists between institutional distance and geographic distance? The contribution provided by the present paper consists in the analysis of a longer period of time - those papers analyse only the publications from 2003-2005- so that it is also possible to observe the evolution of the phenomena in a quite long period, and in the use and comparison of several measurements of institutional distance. This paper shares with the paper by d'Amore, Iorio and Stawinoga (2016) the idea that, regarding the relationship between geographic and institutional

distance, there are both a "resource effect" and a "competence effect". According to the first one, supported by the existing literature, as both kinds of distance imply a cost, when one kind of distance increases, the firm, trying to save resources, reduces the other kind of distance; therefore a trade-off exists. According to the second one, as highly specialised competencies are very dispersed on a geographic and institutional point of view, a firm that needs on-the-frontier knowledge activates an international and heterogeneous network of collaboration; on the contrary, if the required knowledge is more ordinary, a local and homogeneous network is activated; therefore, institutional distance and geographic distance move in the same direction.
In this paper, thanks to a more articulated geographic distinction and a longer considered period, we if such effects have different strengths at different geographic scales and if they vary overtime.

Indeed, the novelty of this study compared to our previous papers is also enrichment compared to the previous literature: the length of the considered period is unusual, as is the use and comparison of different measurements of institutional distance. Above all, the contribution of this paper consists in the more nuanced view of the relationship identified in the existing literature. Other papers supposed that a relationship exists between institutional distance and geographic distance; we analyse whether this connection is constant over time, at different geographic scales, with different forms of research, if it is robust to the change of measurements.

The paper is structured in the following way: the second section presents a review of the more relevant literature on the theme of institutional distance and geographic distance and their relationship in scientific collaborations; the third section presents some hypotheses that may be formulated regarding the relationship between spatial and institutional distance; in the fourth section a description of the data used for the empirical analysis can be found; the fifth session presents the results of the bivariate and multivariate analysis, with a conclusive synthesis of the results; some final considerations conclude the paper.

2. Spatial and Institutional Distance: A Literature Review

The specific focus of this paper is on the relationship between institutional and spatial distance in research collaborations: while an abundant theoretical and empirical literature analyses the effect of spatial distance on R&D collaboration, a more limited number of papers considers how this effect is mediated by other factors, like institutional distance.

A first remark regarding the relationship between institutional distance and geographic distance may be found in Pavitt (1984) and DeSolla Price (1984), who assumed that collaboration between academic and non-academic organisations was more localised into space than collaboration between universities. Boschma (2005) explicitly states that there is an inverse relationship between geographical and institutional proximity, as geographical proximity may compensate for the lack of institutional proximity and institutional proximity facilitates interaction over a long geographic distance. An important reference point for the present study is the paper by Ponds, Van Oort and Frenken (2007), as they analyse the role of geographical proximity for scientific research collaboration in science-based technologies between three kinds of institutions—academic organisations, firms and governmental or non-profit organisations. They observed the co-authorships of scientific publications in eight technological fields as registered in ISI-Web of Science from 1988 onwards. The spatial distance is calculated in great detail: it is the average travel time between the regions (defined at a NUTS-3 level) where the institutions are

located. The consideration of institutional distance is less in detail: as reported above, three kinds of institutions are defined, and there is a distinction between collaboration among institutions of the same kind (no institutional distance) and collaboration among institutions of different kind (institutional distance). Their conclusion is that the collaboration involving different kinds of institutions is more localised than collaboration involving the same kind of institutions: when institutional distance increases, spatial distance reduces.

The paper of McKelvey, Alm and Riccaboni (2003) is of great importance to the present paper too, as the contents of their studies are similar and the sector they analyse is the same (biotechnologies); they also consider a national case (Sweden). They also find a trade-off between spatial and institutional distance: geographical co-location is more important for inter-institutional collaboration (for instance, firms with universities) than for collaboration among the same kind of institutions (for instance, firms with firms and universities with universities).

As reported in the introduction, the paper by d'Amore et al. (2013) is the starting point of the present paper. They analyse a three year dataset of publications in the Italian biotech sector; they identify four kinds of institutions (firms, universities, research centres, and hospitals) to which the authors of the paper belong and provide a measurement of institutional distance, adapting to this purpose the E-I index, a measurement of relative heterogeneity of the collaborations usually used in the social network analysis; they also compare the observed value of this index with a theoretical value deriving from randomisations of the institutions. They find results consistent with the trade-off hypothesis, as the publications deriving from international collaborations are more characterised by a smaller institutional distance than national ones. They are also able to classify the papers according to the stage of the research and they find that the institutional/geographic trade-off holds both for basic and applied research.

Other papers by the same authors analysing the same dataset try to consider other aspects of the relationship between institutional distance and geographic distance. D'Amore, Iorio and Stawinoga (2016) find that the relationship is inverse among rarely cited papers (supposing they are derived from research projects of limited quality or relevance), while the relationship is direct among frequently cited papers (derived from research projects of high quality or relevance). They also find that different institutions have different behaviours and attitudes towards managing the balance between the two kinds of distance.

D'Amore at al. (2012) find that a trade-off between the two kinds of distance may be identified only in occasional collaborations, while a relationship of complimentarity (more spatial distance, more institutional heterogeneity) holds, when the relationships are stable.

Other papers, even though not exactly focused on the institutional/geographic trade-off, are relevant to this work, as they analyse the effect of spatial distance on collaborations through the interaction of other factors, first of all the kind of research that is one of the variables included in this analysis.

Broström (2010) explores whether in university-industry interactions there is a relationship between the spatial distance and the kind of research involved in the project. He conducts a survey among the managers responsible for R&D in the engineering sector in Sweden and he finds that geographical proximity is important for short-term projects of a very applied nature because the exchange of tacit knowledge is particularly relevant for this kind of research, while in long-term projects it is generally easier to work across geographical distance.

Muscio (2012), who conducted almost two hundred interviews in Italian university departments, finds a different result: the applicability to industry of research has a positive impact on the probability of establishing collaborations with distant firms. He also studies other determinants of the distance of collaborations, finding that size of department has a negative effect on the distance of collaborations and that there is a sort of complementarity among collaborations at different geographic scales; for example, for a university department, having been involved in regional or national collaborations increases the probability of collaborating with European and extra-European firms.

D'Este and Iammarino (2010), studying collaborating research in engineering and the physical sciences in the UK, find that in university-industry partnerships, high quality of the university departments increases the probability of attracting distant business partners, but this only holds until a certain threshold of research excellence: beyond such threshold, collaboration with industry tends to be geographically closer.

Autant-Bernard et al. (2007) analyse the role of geographical distance and of the "network effects" (the position and role in the network of collaboration) in affecting the probability of collaborating in R&D projects. The analysis is conducted among participants in 290 research projects submitted for the 6th EU Framework Programme on micro and nanotechnogies. They distinguish among firms that are involved in many projects and firms involved in one single project. Taking into consideration only the "multi-project" firms, there is no evident influence of spatial distance on the probability of collaborating, while there is a clear influence of the firm's position within the network (number of direct and indirect partners, social distance between firms). If "single-project" firms are taken into account too, both geographical distance and social network effects matter, reinforcing the phenomenon of intra-national local clustering.

Even Scherngell and Barber (2009) find different effects of geographical distance on R&D collaboration in two different groups of agents: in this case the distinction is among private and public agents. Considering the collaboration among firms (industrial R&D networks), spatial distance seems to have an important effect on the probability to collaborate, while, analysing the public research R&D network among universities and research organisations, the effects of geography are smaller. In both groups the technological distance is the most important factor. This analysis is conducted among the projects of the 5th EU Framework Programme.

The same authors, analysing the same data, find that the spatial proximity does increase the probability of collaborating between different organisations, but other factors may act in the same way; such factors include the thematic proximity, the experience in projects of the same kind, the prior acquaintance, and the centrality of the institution in the network of collaborations (Scherngell and Barber, 2011).

3. Expectations about the Relationship between Spatial and Institutional Distance

The existing literature about the relationship between spatial and institutional distance in the research collaborations, reviewed in the previous section, suggests some hypotheses. Following Ponds, Van Oort and Frenken (2007), it is possible to argue that both kinds of distance, spatial and institutional, imply a cost that may be intended in direct monetary terms or in terms of "strength". As agents try to minimise costs, there is a trade-off between the two kinds of distance:

the more the spatial distance, the less the institutional distance and *vice versa*. Therefore, local networks should be more heterogeneous than international networks. This may be called a "resource based" effect.

Indeed, another line of argument is possible, leading to an opposite conclusion: highly specialised competencies are very dispersed on a geographic and institutional point of view and a firm that needs on-the-frontier knowledge activates an international and heterogeneous network of collaboration; if, on the contrary, the required knowledge is more ordinary, a local and homogeneous network may be activated. If this argument is correct, local networks should be more homogeneous than international networks. This may be called a "competence based" effect).

The first effect is substantially based on an idea of at least partial substitutability of collaborators (considered like a sort of inputs in the production of new knowledge) therefore, the costs of factors have a role in the choice process. The second effect is based on the idea that particularly high competencies are not easily substitutable and must be deployed "wherever they are".

It is, of course, possible, if not likely, that the two effects coexist. This may generate ambiguous results that are not a clear direction of the relationship between institutional distance and geographic distance. It is also likely that the two effects have different strengths at different geographic levels—it is, for instance, likely that the importance of the resource-based effect significantly increases at a large geographic scale when the costs generated by the geographic distance become very high. The most plausible expectation is, therefore, that the direction and strength of the relationship vary with the geographic scale. Besides, it is likely that the effects may vary their intensity over time.

The studies we cited do not properly address this kind of problems: the models used for statistical analysis usually do not allow for a complex relationship between the two kinds of distance, neither the evolution over time is considered. Besides, the existing literatures leave the problem of measurement of institutional distance open: each paper adopts one single measurement of institutional distance, but different papers adopt different measurements so that there is no consensus about the "best" measurement, neither is the stability of the results tested with the variation of measurements. We try to overcome this limitation, analysing the same data with different measurements.

4. Data

In order to build a database of scientific publications in the biotech sector, the biotech sector has been made an intersection of two databases: *i*) RP Biotech data base; *ii*) ISI Web of Science. They are briefly described as follows.

RP Biotech data base. It is a collection of potentially all the Italian firms belonging to the biotech sector, active as of December 2005. In this study only the 306 life-science for profit firms are considered.

ISI databases, especially the Science Citation Index®, and the web-based version Web of Science, is a detailed bibliometric database of journal articles and citations of worldwide research literature that contains 14,000 international peer-reviewed scientific and technical journals.

There is information about publications of the selected firms across the period 1990-2005. The record of each publication in the ISI-Web of knowledge reports contains, among other kinds of information, the name of the authors and the name of the institutions to which they belong. All the publications where the name of at least one of the selected firms (Italian life-science for-profit biotech firms) appeared among the institutions of affiliation were extracted. Then four categories of institutions were identified (firms, universities, research centres and hospitals) and it was established to which category each institution belongs.[1]

Besides, the papers were classified according the nature and the stage of the research -whether basic or applied- and according to the criterion suggested by Lewison and Paraje (2004).

5. Empirical Results

5.1 Measurements of Institutional Distance and Bivariate Analysis

In the empirical analysis, two measurements are extensively used:

- The number of institutions to which the authors of each publication belong (we call it NUMINST);
- The number of kinds of institutions to which the authors of each publication belong (we call it KINDINST).

An example may be useful to clarify their meanings. If a paper has been written by authors belonging to the University of Milan, the University of Rome, the Hospital of Turin, the firm Rossi S.P.A and the firm Verdi S.P.A., the number of institutions is five (NUMINST = 5) and there are three different kinds of institutions: universities, hospitals and firms (KINDINST = 3).

Table 1 reports some relevant statistics about the sample of Italian biotech firms' publications: there is the total number of publications, the number and percentage of publications deriving from institutional collaboration, the mean value of NUMINST and the mean value of KINDINST. The table has four columns, the first one reporting the values for the whole considered period (1990-2005), the other for three sub-periods (1990-95; 1996-2000; 2001-05). It is possible to observe that the institutional collaboration is an increasing phenomenon: from the first period (1990-95) to the third period (2001-2005), the percentage of papers written with the collaboration of at least two institutions increases from 78.4% to 86%; the mean number of institutions increases from 2.59 to 3.37; the variety of involved institutions increases too, as the mean value of KINDINST goes from 1.94 to 2.16.

Table 1: Descriptive statistics about publications and involved institutions (all the papers)

Period	90-05	90-95	96-00	01-05
Publications	3863	1134	1095	1634
Publications in collaboration	3205 (83%)	889 (78.4%)	910 (83.1%)	1406 (86%)
Number of institutions per paper (NUMINST) – mean value	3.09	2.59	3.19	3.37
Number of kinds of institutions per paper (KINDINST) – mean value	2.07	1.94	2.06	2.16

In order to analyse the impact of spatial distance, the publications were divided into three categories: national regional publications (those written by authors all belonging to institutions located in the same Italian region); national extra-regional publications (those with all the authors belonging to Italian institutions, with at least two institutions located in two different Italian regions); international publications (those with at least one author belonging to a non-Italian institution).

In the following analysis, publications not written in collaboration (those written by authors belonging only to an Italian biotech firm) were excluded; therefore, the analysis is done on 3205 publications. Among these, 620 (19.3%) are national regional publications, 1399 (43.6%) are national extra-regional publications and 1186 (37%) are international publications.[2]

Table 2 reports the geographical nature of the publications in the whole period and in the three sub-periods. The rate of international papers increases from 31.5% in the first period (1990-95) to 38.8% in the third period (2000-05).

Table 2: Geographical nature of the publications

	1990-2005	1990-1995	1996-2000	2001-2005
Regional publications	620 (19.3%)	202 (22.7%)	188 (20.7%)	230 (19.4%)
Extra-regional publications	1399 (43.6%)	407 (45.8%)	361 (39.7%)	631 (44.9%)
International publications	1186 (37%)	280 (31.5%)	361 (39.7%)	545(38.8%)
All publications	3205	899	910	1406

There is not a universally accepted measurement of institutional distance: it is commonly accepted that a collaboration between two institutions of the same kind implies less institutional distance than a collaboration between two different kinds of institutions, but, except for the paper by d'Amore et al. (2013), no measurements that go beyond such consideration have been

suggested. Starting from the commonly accepted statement reported before, it was assumed that an increase in the variety of institutions increases the institutional distance; a co-authorship between authors belonging to one university, one hospital or one firm implies more institutional distance than a co-authorship between authors belonging to two universities and one firm. Therefore, KINDINST, which measures exactly how many kinds of institutions are involved in a co-authorship, may provide a good measurement of institutional distance. As four kinds of institutions were identified, this index ranges from 1 to 4.

KINDINST is an absolute measurement of institutional distance, as it is not normalised for the total number of institutions. But, if the total number of institutions involved in a paper increases, the expected value of this index increases too. As collaborations on larger geographic scales have a higher probability of including a higher number of institutions (an international paper may include national and foreign institutions, and a national paper may include only national institutions) international papers will have a higher expected value of KINDINST than national papers and extra-regional than regional, being choices and preferences of the agents constant across different geographic scales. In other words, a random combination of institutions will generate a higher expected value of KINDINST at larger geographic scales.

Table 3 and Table 4 report respectively the value of NUMINST and KINDINST for each geographic scale. In the first column the values for the whole period are reported; in the second, third and fourth columns there are the values for the three sub-periods. The observed values of NUMINST across geographical dimensions of publications are in line with the expected results, as the mean values increase when geographical scale becomes wider. KINDINST significantly increases between regional and extra-regional levels and slightly decreases between extra-regional and international level: as stated above, the first result may simply derive from a random combination of the institutions; the second result, as it moves in different direction from random expectations, may be interpreted as a preference of the agents toward less institutional distance when the geographic scale becomes wider.

Table 3: Mean values of NUMINST

	1990-2005	1990-1995	1996-2000	2001-2005
Regional publications	2.60	2.43	2.63	2.72
Extra-regional publications	3.68	3.14	4.12	3.77
International publications	3.82	3.32	3.67	4.17
All publications	3.52	3.03	3.63	3.75

Table 4: Mean values of KINDINST

	1990-2005	1990-1995	1996-2000	2001-2005
Regional publications	2.13	2.11	2.11	2.16
Extra-regional publications	2.34	2.25	2.34	2.40
International publications	2.30	2.20	2.29	2.36
All publications	2.29	2.20	2.27	2.35

In order to obtain more robust results, it is necessary to obtain a measurement of institutional distance that "controls" for the increase in the number of institutions. In D'AMORE et al. (2013) this problem is overcome by comparing the observed institutional variety with the expected value generated by random combinations of the institutions. In this paper some different measurements

of institutional distance that control for the number of institutions are provided but using exclusively observed values.

A simple way to control for the variation in the number of institutions is to observe what happens to the institutional variety if the number of institutions involved in a publication is given. Considering all the papers written by authors belonging to two institutions (when NUMINST is equal to two) the mean value of KINDINST at each geographic scale was calculated; the same was done for all the papers written by authors belonging to three, four or five institutions (when NUMINST is equal to three, four, five).

Table 5 reports the values of KINDINST on the rows for regional, extra-regional, and international papers, then for all the papers; on the columns for NUMINST equal to two, three, four and five. The number of papers belonging to each group is reported in brackets. As it was easy to forecast, the institutional variety increases when the number of involved institutions increases: among papers with NUMINST equal to two, the mean value of KINDINST is 1.87; the mean value of KINDINST is 2.29, 2.50, 2.79 respectively among papers written by authors belonging to three, four and five institutions. The hypothesis of a trade-off between institutional distance and geographic distance would be supported by a decrease of the value of KINDINST when the geographic scale enlarges for each given NUMINST. This happens with certain evidence between extra-regional and international scales when NUMINST is equal to two and three the value of KINDINST decreases from 1.92 to 1.71 and from 2.38 to 2.18. In such cases the differences between the means is significant (at 95% level) according to a standard t-test; other differences of the mean values are not significant at 95% level. It must be observed that these two groups of papers -with NUMINST equal to two and three- include two-thirds (2096 over 3205) of all the considered papers.

Table 5: Mean values of KINDINST for different numbers of institutions 1990-2005

	NUMINST=2	**NUMINST=3**	**NUMINST=4**	**NUMINST=5**
Regional publications	1.94 (375)	2.30 (160)	2.48 (58)	2.89 (19)
Extra-regional publications	1.92 (479)	2.38 (412)	2.50 (236)	2.76 (114)
International publications	1.71 (317)	2.18 (353)	2.50 (215)	2.80 (147)
All publications	1.87 (1171)	2.29 (925)	2.50 (509)	2.79 (280)

If this analysis is repeated for the three sub-periods considered before, results are absolutely similar in each of them: the difference in the value of KINDINST is significant between the international and extra-regional papers when NUMINST is equal to two and three.[3]

Another way to control for the expected variation of number of institutions across the geographic scale consists in calculating some relative indexes of institutional distance, normalised for the total number of institutions or for the total number of "linkages" among institutions.

A relative measurement of institutional distance derives, in its definition and terminology, from the social network analysis (Krackhardt and Stern, 2008; d'Amore et al., 2013). We considered the institutions involved in the publications as "nodes" of a network; if, for instance, authors belonging to three institutions are co-authors of a paper, there is a linkage between those institutions; that linkage may be homogeneous or "internal" if it happens between two institutions of the same kind (e.g. two universities, two firms, etc.) heterogeneous or "external" if it happens between two institutions of different kinds (e.g. one university and one firm, one firm and one hospital, etc.) For each paper the number of "external" linkages divided for the total number of linkages ("external" plus "internal" ones) is a good relative measurement of institutional distance. We call this index the *E index.*[4] We have, therefore:

E index = number of external linkages / (number of external linkages + number of internal linkages)

This index theoretically ranges from 0, when all the linkages are internal (all the institutions are of the same kind) to 1, when all linkages are external (all the institutions are of a different kind). This index has no meaning when there is not a collaboration, that is, if a paper has been written by authors belonging to only one institution: in this case there are no linkages and the index is equal to 0/0.

Table 6 reports in the first three rows the mean value of the *E index* for the different geographic levels of the publications and in the fourth row the mean value for all the papers; the first column reports the values for all the period and the other three columns for the three sub-periods. It may be observed that in each sub-period and in the whole considered period, the value of the *E-index* decreases -therefore, the institutional distance decreases- when the geographic scale enlarges. This result is consistent with the presence of the trade-off between institutional distance and geographic distance.

Table 6: Mean values of *E index*

	1990-2005	**1990-1995**	**1996-2000**	**2001-2005**
Regional publications	0.85	0.91	0.82	0.83
Extra-regional publications	0.78	0.81	0.78	0.76
International publications	0.68	0.71	0.69	0.66
All publications	0.76	0.80	0.75	0.73

The *E index* may be calculated also at an institutional level: it is possible to calculate how many "internal" and "external" linkages the universities, the hospitals, the research centres and the firms have in each paper than to calculate an average value for each kind of institution. This sample is fully representative for firms, as it is built on firms' publications (it may be useful to remind that there are all the publications done by the Italian biotech firms that were active in 2005), therefore, it is particularly interesting to calculate the *E index for firms*. Table 7 reports the value of this index for the different geographic levels and for the three sub-periods. Between the extra-regional and international levels, the index decreases in each sub-period, while, if regional and extra-regional level are compared, the index decreases in the period 1996-2000 and remains constant in the period 2001-2005. The behaviour of the firms seems, therefore, less clearly oriented towards an institutional/geographic trade-off than the behaviour of the four kinds of institutions globally considered.

Table 7: Mean values of *E index for firms*

	1990-2005	1990-1995	1996-2000	2001-2005
Regional publications	0.95	0.98	0.92	0.94
Extra-regional publications	0.94	0.95	0.95	0.94
International publications	0.83	0.86	0.82	0.82
All publications	0.90	0.93	0.89	0.89

5.2 Multivariate Analysis

A further step in this analysis is represented by the regression analysis, which lets us overcome the intrinsic limits of the analyses reported above and to take other factors into consideration. The limit of the analysis with KINDINST keeping NUMINST constant consists in the lack of a synthetic view: the comparison between geographic levels may be done for each given number of NUMINST, but it is not possible to synthesize in a unique number the presence or absence of the institutional or geographic trade off. On the other side, the expected value of the *E index* is not independent of the number of involved institutions; therefore, its value, at different geographic scales, may depend not only on the institutional variety -that is what it is desired to measure-but also on the number of involved institutions, which, as reported before, has an increasing trend with the enlargement of geographic scale. A regression model may solve both problems by having a measurement of institutional distance as an dependent variable, a measurement of geographic distance as another independent variable, and introducing, as control variables, as many dummy variables as each value of NUMINST in such a way a restriction is imposed (the relationship between geographic and institutional distance is the same whatever the number of involved institutions) but a synthetic measurement of the relationship may be obtained, controlling for the number of institutions.

In addition to controlling for the number of institutions, the regression analysis lets us take into account other factors that may influence the relationship between institutional distance and geographic distance. First of all, the nature or the stage of the research, if basic or applied, may systematically affect the structure of the collaborations and, thereby, their institutional and geographic nature. As illustrated in Iorio et al. (2012), where the same dataset of this paper is analysed, all the papers of the dataset are classified adapting a criterion suggested by Lewison and Paraje (2004), using the word in the titles as filters. Considering only the publications written in collaboration, it has been possible to classify 3146 out of 3205 publications: 1295 (38%) papers are basic, 1555 (49.4%) are applied and 396 (12.6%) are mixed between basic and applied. Then, the introduction of a time dummy variable may enable us to control for temporal trends. Besides, as the database is built on the biotech firms, it is possible that different firms have different needs or policies regarding the institutional distance and the geographical distance of their collaborations and their publications. This may be controlled by the introduction of firm dummy variables.

The estimated models are, therefore, so structured:

- The <u>unit of analysis</u> is the single publication;
- The <u>dependent variable</u> is a measurement of institutional distance (KINDINST, *E index* or *E index for firms*);
- The <u>determinant independent variable</u> of interest is the geographic level of the publications; it is expressed by two dummy variables assuming value 1 if the publication belongs to that level and value 0 if it does not (REGIO assumes value 1

if the publication is regional, 0 otherwise; INTERNAT assumes value 1 if the publication is international, 0 otherwise; regional and international papers are, therefore, compared with extra-regional papers);

- The control independent variables are the number of institutions to which the authors belong, the kind of research of the publications, the time variables and the firm variables. More in detail:

  - The number of institutions is expressed by four dummy variables NUMINST_*n* with *n* equal to 3, 4, 5 and >5, assuming value 1 if the publication has been written by authors belonging to *n* institutions and 0 otherwise (the variable NUMINST_3 has value 1 if the paper has been written by authors belonging to three institutions and 0 otherwise; NUMINST_4 has value 1 if the paper has been written by authors belonging to four institutions and 0 otherwise; NUMINST_5 has value 1 if the paper has been written by authors belonging to five institutions and 0 otherwise; in order to limit the number of variables and to make the results more easily reportable, for NUMINST greater than 5 we introduce a unique dummy variable, NUMINST_>5, which has value 1 if the paper has been written by authors belonging to more than five institutions and value 0 otherwise; the benchmark variable is, therefore, represented by the papers written by authors belonging to two institutions).

  - The nature of the research is expressed by a dummy variable, indicating whether the paper is of basic nature or not: BASIC assumes value 1 if the paper is of basic nature, 0 otherwise, that is whether it is applied or mixed; basic papers are, therefore, compared with non-basic papers.

  - Temporal trends are controlled by two dummy variables for the second and third last sub-period (*Years 1996-2000* assumes value 1 if the paper has been published in one of those years, 0 otherwise; *Years 2001-05* assumes value 1 if the paper has been published in one of those years, 0 otherwise; therefore, papers published in those two periods are compared with papers published in the period 1990-1995)

  - Firm behaviour is controlled by a dummy variable for each Italian biotech firm that co-authored a publication.

Besides a complete model, including all these variables, two "restricted" models are also estimated: Model 1, excluding the dummy variables for nature of research, time and firms, and Model 2, excluding firm dummy variables; the complete model is, therefore, called Model 3. Moreover, separated estimations for each of the three sub-periods are done.

When KINDINST is the dependent variable, as it may only assume integer values from 1 to 4, an ordered probit is the suitable model; when the *E index* and the *E index for firms* are the dependent variables, as they assume (almost) continuous values, a linear regression analysis, estimated with ordinary least squares (OLS), is the correct model to estimate; because of the presence of heteroskedasticity of errors, we estimated regressions with robust standard errors.

Table 8 reports the results of the ordered probit analysis having KINDINST as a dependent variable; the three columns report the results respectively for Model 1, Model 2 and Model 3. The most interesting result, for the aim of this paper, is that international publications have *ceteris paribus* a significantly (at 99% level) lower value of KINDINST than extra-regional papers, while the difference between extra-regional and regional papers is not statistically significant. It is, therefore, possible to conclude for the existence of a trade-off between institutional distance and geographic distance if international and national levels are compared, while there is no statistically robust confirmation of the hypothesis of the trade-off if regional and extra-regional papers are compared. This result is analogous to what d'Amore et al. (2013) found analysing the same data but for a more limited period of time with a different measurement of institutional distance and estimating a similar, but not identical, model. These results are consistent with the hypothesis of the co-existence of two effects -the resource based and the competence based- which may have different strengths at different geographic levels. When the physical distance among the institutions is high, the costs of physically distant collaborations become so high that there is need to save resources on the institutional side and the resource based effect, which implies a trade-off between the two kinds of distance, prevails. At a smaller geographic scale, the need to "save money" is not so strong; therefore, the two effects are of similar intensity, generating an ambiguous result.

Easily predictable, the signs of the dummy variables NUMINST_*n* show that institutional variety expressed by KINDINST increases when the number of institutions involved in the publication increases. Comparing Model 1 and Model 2, it may be observed that the inclusion of dummy variables for the nature of research and for sub-periods only slightly increases the goodness of fit of the model but has an effect on the strength of the institutional or geographic trade-off, making it stronger (the coefficient for international papers is much higher in Model 2 than in Model 1). The nature of research is significantly (at 99%) related with the institutional distance: basic research papers are characterised, on average, by a higher institutional variety than applied and mixed papers. The inclusion of firm dummy variables (Model 3) more significantly increases the goodness of fit of the model, indicating that firms do have different behaviour in terms of institutional variety of collaborations. Besides, the inclusion of such dummy variables has some interesting consequences for the other coefficients: the magnitude of the coefficients of the dummy variable for geographic dimensions of papers returns very close to those of Model 1 and the coefficient of the dummy variable for period 2000-2005 becomes positive and significant (the coefficient for the period 1995-2000 becomes positive too, even if not significant)[5]. These results mean that when the complete model is estimated, controlling, therefore, for the number of institutions, the nature of research, time trends and firm specificity, the trade-off between institutional distance and geographic distance is confirmed if national and international levels are compared. The institutional variety increases when the number of involved institutions increases. Basic research papers are characterised by a greater institutional variety than applied papers and more recent years are characterised by a higher institutional variety than the early years of the considered period.

Table 9 reports the results of the estimation of the complete model, with KINDINST as dependent variable, for each of the three considered periods (time dummy variables are excluded, as they are constant inside each period). The decrease in institutional distance, moving from extra-regional to international papers, is observed in each of the three sub-periods, but its magnitude is much stronger in the first sub-period: the institutional/geographic trade-off persists but has reduced its intensity, perhaps because the decrease in travel costs lowered the costs of physical distance, reducing the resource based effect. The more evident

difference among the three sub-periods concerns the sign and significance of the coefficient of the dummy variables for basic research papers: the relationship between the basic nature of research and the institutional variety, which in the previous analysis was found positive and significant for the whole period, indeed progressively increases and becomes positive and significant only in the more recent period.

Table 8: Results of the ordered probit-determinants of KINDINST (years 1990-2005)

	Dependent variable: KINDINST (ordered probit) Years 1990-2005		
	Model 1	**Model 2**	**Model 3**
Independent Variables	**Coefficients (p-value)**	**Coefficients (p-value)**	**Coefficients (p-value)**
Regional publication	0.042 (0.494)	0.034 (0.611)	0.045 (0.520)
International publication	-0.283*** (0.000)	-0.584*** (0.000)	-0.278*** (0.000)
Benchmark: extra-regional publication			
NUMINST_3	1.157*** (0.000)	1.142*** (0.000)	1.247*** (0.000)
NUMINST_4	1.622*** (0.000)	1.608*** (0.000)	1.745*** (0.000)
NUMINST_5	2.206*** (0.000)	2.200*** (0.000)	2.377*** (0.000)
NUMINST>5	2.646*** (0.000)	2.642*** (0.000)	2.801*** (0.000)
Benchmark: NUMINST_2			
Basic research	-	0.116*** (0.008)	0.169*** (0.001)
Bechmark: applied and mixed research			
Years 1996-2000	-	-0.041 (0.294)	0.100 (0.139)
Years 2001-2005	-	-0.015 (0.777)	0.167** (0.016)
Benchmatk: Years 1990-2005			
Firm dummy variables	Excluded	Excluded	Included
Statistics			
Number of observations	3204	3145	3145
Log likelihood	-2425.588	-2380.9849	-2166.0647
LR Chi2 (df)	1359.10 (6)	1336.91(9)	1766.75 (153)
Prob>Chi2	0.000	0.000	0.000
Pseudo R^2 (McFadden)	0.2188	0.2192	0.2897

***Significant at 99%; **Significant at 95%

Table 9: Results of the ordered probit-determinants of KINDINST (three periods)

	Dependent variable: KINDINST (ordered probit)		
	Model 3 Years 1990-1995	**Model 3 Years 1996-2000**	**Model 3 Years 2001-2005**
Independent Variables	**Coefficients (p-value)**	**Coefficients (p-value)**	**Coefficients (p-value)**
Regional publication	-0.002 (0.990)	0.037 (0.786)	0.078 (0.499)
International publication	-0.351*** (0.001)	-0.241** (0.026)	-0.245*** (0.003)
Benchmark: extra-regional publication			
NUMINST_3	1.301*** (0.000)	1.234*** (0.000)	1.402*** (0.000)
NUMINST_4	1.940*** (0.000)	1.863*** (0.000)	1.775*** (0.000)
NUMINST_5	2.434*** (0.000)	2.689*** (0.000)	2.472*** (0.000)
NUMINST>5	2.510*** (0.000)	2.866*** (0.000)	2.998*** (0.000)
Benchmark: NUMINST_2			
Basic research	-0.018 (0.858)	0.167 (0.135)	0.300*** (0.000)
Bechmark: applied and mixed research			
Firm dummy variables	Included	Included	Included
Statistics			
Number of observations	872	895	1378
Log likelihood	-520.2075	-579.6975	-971.207
LR Chi2 (df)	373.20 (50)	599.05(94)	920.27 (135)
Prob>Chi2	0.000	0.000	0.000
Pseudo R^2 (McFadden)	0.2640	0.3407	0.3215

***Significant at 99%; **Significant at 95%

Table 10: Results and statistics of the linear regression of determinants of *E index* and *E index for firms* (Model 3-Years 1990-2005)

	Dependent variable: *E index* (linear regression with robust standard errors)	**Dependent variable: *E index for firms* (linear regression with robust standard errors)**
Independent Variables	**Coefficients (p-value)**	**Coefficients (p-value)**
Regional publication	0.015 (0.234)	0.017 (0.146)
International publication	-0.074*** (0.000)	-0.108*** (0.000)
Benchmark: extra-regional publication		
NUMINST_3	-0.128*** (0.000)	0.037*** (0.002)
NUMINST_4	-0.205*** (0.000)	0.058*** (0.000)
NUMINST_5	-0.199*** (0.000)	0.059*** (0.000)
NUMINST>5	-0.267*** (0.000)	0.111*** (0.000)
Benchmark: NUMINST_2		
Basic research	0.032*** (0.002)	0.034*** (0.001)
Benchmark: mixed and applied research		
Years 1996-2000	0.009 (0.497)	0.005 (0.690)
Years 2001-2005	0.021 (0.127)	0.006** (0.655)
Benchmark: Years 1990-1995		
Firm dummy variables	Included	Included
Number of observations	3145	3145
Statistics		
F (df)	7.69 (154, 2990)	6.24 (120, 2990)
Prob>F	0.000	0.000
Adjusted R^2	0.244	0.2432

***Significant at 99%; **Significant at 95%

Table 10 reports the results of the estimation of the complete model for the whole period, having the *E index* and the *E index for firms* as dependent variables. The results of both estimations are similar to the complete model estimated with KINDINST as dependent variable, confirming the trade-off between international and extra-regional levels and the

greater institutional variety of basic publications. The only relevant difference concerns the effect of dummy variables for sub-periods that are not significant when the *E index* is the dependent variable.

5.3 Summary of Results

The more relevant results of the empirical analysis may be synthesised in the following way:

- There is an increase over time in the number and variety of institutions involved in publications.

- When geographical scale enlarges, the number of institutions involved in publications enlarges; this may depend in part on the greatest number of existing institutions if a larger geographical scale is considered.

- In a bivariate analysis, using KINDINST as a measurement of institutional distance and controlling for the number of institutions, it is possible to find a trade-off between institutional distance and geographic distance if international and national papers are compared and the number of involved institutions is less than four.

- In a bivariate analysis, using *E index* and *E index for firms* as measurements of institutional distance, the trade-off between institutional distance and geographic distance is found whatever the geographic level compared.

- In a multivariate analysis, using KINDINST, *E index* and *E index for firms* as measurements of institutional distance, controlling for several factors (the number of institutions involved in each publication, the nature of the research, the time trends, and the individual firm behaviour) and conducting separated analyses for three different sub-periods, the following results emerge:

    - It is possible to find a trade-off between institutional distance and geographic distance if international and national papers are compared.
    - The trade-off between institutional distance and geographic distance was more intense in the early period (first five years) than in the latter years.
    - Basic research publications are characterised by a greater institutional distance than applied and mixed (basic/applied) ones.
    - When the analysis is conducted in the three sub-periods, the previous effect, increasing over time, is statistically significant only for the most recent publications (last five years of the considered period).
    - In recent years, there was an increase in the institutional variety (but when the *E index* is the dependent variable, this effect is not significant).

6. Conclusions

In this paper the co-authoring networks in the scientific publications of the firms belonging to the Italian biotech sector are analysed in order to understand the relationship existing between geographic distance and institutional distance.

The existing literature, including a previous study on the same database but for a more limited time period and with different measurements, suggests the existence of an inverse relationship between the two kinds of distance. According to a "resource based" theory, as both kinds of distance are a cost, a trade-off between them exists; therefore, when spatial distance increases, institutional distance decreases. It is suggested that a "competence based" theory is also possible. If highly specialised competencies are very much dispersed on a geographic and institutional point of view, a firm that needs on-the-frontier knowledge collaborates with international and heterogeneous partners; if the required knowledge is more ordinary, collaborations happen with local and homogeneous partners. According to this view an increase in spatial distance goes together with an increase in institutional distance.

Besides the theoretical problem, this paper focuses on the measurement issue: as institutional distance has no universally accepted definition and measurement, several ways to make a measurement of it are suggested and such different measurements are used to empirically investigate our research question.

Using such different measurements of institutional distance, both bivariate and multivariate analyses are conducted, controlling for several factors that may influence the institutional distance: nature of the research, time trends, firm specificity. The empirical analysis finds an inverse relationship between institutional distance and geographic distance, but the different analyses converge to this result only beyond certain geographic scale -if international and national papers are compared- while comparing national papers at different regional and extra-regional geographic levels, the direction of the relationship is not statistically clear, suggesting the possible contemporary presence of an opposite effect, that is, the competence based effect. If the presence of these two effects is true, the fact that the strength of the trade-off is found to be reducing over time could be explained by an increasing importance of the competence-based effect.

These results are, therefore, consistent with the hypothesis, formulated in the theoretical part of the paper, that the two effects coexist, having different strengths and different geographic levels. We also supposed that these effects may change their intensity over time: this hypothesis is confirmed by the change, in the different sub-periods considered, of the intensity of the relationship between spatial and institutional distance.

Another interesting result, little explored in previous literature, concerns the nature of the research and its increasing effect on the institutional variety. In recent years, publications with basic research content showed a higher institutional variety than more applied papers.

In summary, the analysis presented in this paper reconsiders topics covered in other papers, giving them a more rich and nuanced view, taking more aspects into consideration: the different geographic scales, the long period of time, the nature of research, the variety of measurements.

Anyway, it must be added that this empirical analysis has some limits, as there is not a fine measurement of geographic distance and several important variables (source of funds, with possible different incentives to institutional collaboration, dimensions of the institutions, etc.) are not taken into consideration[6]. Indeed, this study should be considered as a further step in the way indicated by previous analyses: the convergence of results obtained analysing different kinds of collaborations, different sectors and countries, using different measurements

and techniques, show that an initial intuition about the relationship between different kinds of distance is becoming a consolidated point. A further and more in-depth analysis that takes other elements into consideration now need to be conducted on this topic, as it may shed more light on the way knowledge flows in an innovative sector and highlight what should be taken into consideration by the policy maker that aims to promote research collaborations between different institutions. A careful consideration of the role played by different kinds of distance in managing the proximity issue may be useful in designing proper policy measurements in a field that is so important for competitiveness in the contemporary knowledge-based economy.

Endnotes

[1] A more detailed description of the Italian biotech sector may be found in d'Amore and Vittoria (2008, 2009) and in d'Amore, Iorio and Stawinoga (2010).

[2] It is important to underline that in this paper, there is an analysis of the relationship between the institutional distance and the geographic scale, which is something different and richer than the pure physical distance. In fact, when geographic scale enlarges, it is not only the physical distance among the authors and the institutions that increases, but differences in norms, culture and languages increase too. All such differences represent barriers to overcome with "strength", therefore, a cost. Regarding physical distance, it is, of course, possible that, in some cases, an enlargement of geographic scale does not correspond to an increase in physical distance (the distance between two towns both located in Italy but in two different regions could be greater than that between two towns located near the national borders, but with one in Italy the other abroad) but, in the data analysed in this paper, this relationship is on average largely verified. Regarding the other barriers, inside the same region there are the same norms, the same regional policy and a common innovative and cultural *milieu*, with possible clustering phenomena; across Italian regions the regional laws may differ but actors share the same national legislative framework and the same national policy. International collaborations involve different norms —except common Communitarian laws if partners belong to the European Union— and cultures. Regarding the language, there are no barriers in national collaborations. While language may represent a barrier in international collaboration, in Italy only a few people from small minority communities do not have Italian as their main language, while, except for a little region in Switzerland, in no other country Italian is spoken as a first language by any community.

[3] The results of this analysis are not reported here.

[4] Krackhardt and Stern (1998) and d'Amore et al. (2013) use the *E-I index,* that is equal to*:* (number of external linkages - number of internal linkages) / (number of external linkages + number of internal linkages); anyway the *E index* and the *E-I index* are perfectly positively correlated; therefore, using one or another leads to the same results.

[5] We also tested the models with a dummy variable for each year: the results are very similar to the models with dummy variables for the three sub-periods.

[6] Other variables not included in the analysis, like dimension and quality of institutions, may influence both geographic distance and institutional distance and, therefore, their relationship. Public funding programs is certainly a variable not included in our analysis that may affect our result: national and international collaborations may benefit from different public funding programs that may imply different incentives in terms of institutional collaborations. In the case of Framework Programme projects, there is a specific incentive to heterogeneous and international collaborations (indeed, this would be consistent with our hypothesis that important projects involve high institutional distance and geographic distance, even though the reason should be found in an "external" incentive rather than in an "internal" one, as supposed in our analysis). Such data are currently missing and should be collected for the further stages of our analysis.

Attributions and Acknowledgments

This paper comes from a strict collaboration between the authors, both on the conceptual and empirical point of view. Anyway, sections 2, 4 and 5.1 have been written by Rosamaria d'Amore; sections 3 and 5.2 by Roberto Iorio; introduction, section 5.3 and conclusions have been jointly written. The authors are grateful to Pietro Panzarasa for his precious suggestions and encouragement.

References

Autant-Bernard, C, Billand, P, Frachisse, D & 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*, vol.86, no.3, pp. 495-520

Bonaccorsi, A, & Thoma, G 2007, 'Institutional complementarity and inventive performance in nanotechnology', *Research Policy*, vol. 36, no. 6, pp. 251-283.

Boschma, R 2005, Proximity and Innovation: a Critical Assessment, *Regional Studies 39*, 61-74.

Broström, A 2010, 'Working with Distant Researchers-Distance and Content in university-industry interaction', *Research Policy* vol. 39, no. 10, pp.1311-1320.

d'Amore, R, Iorio, R, Labory, S & Stawinoga, A 2013, 'Research collaboration networks in biotechnology: exploring the trade-off between institutional and geographic distance', *Industry and Innovation*, vol.20, no.3, pp. 261-276.

d'Amore, R, Iorio, R, & Stawinoga, A 2010, 'The Role of Institutions and Firm Characteristics in the Network of Firm Publications. An analysis of the Italian Biotech Sector through the Social Network Analysis', *World Journal of Management,* vol. 2, no. 3, pp. 81-100.

d'Amore, R, Iorio, R, & Stawinoga, A 2012, 'How do the institutions involved in scientific collaboration deal with spatial and institutional distance?' An analysis of the co-autorships of scientific publications, *University of Salerno DISES Working paper,* no. 3.322.

d'Amore, R, Iorio, R, & Stawinoga, A 2016, 'The Effect of Research Quality and Different Institutional Behaviour on the Relationship between Institutional and Geographic Distance in Scientific Publications', *World Journal of Social Sciences,* vol.6, no.1, pp. 1-19.

d'Amore, R & Vittoria P 2008, 'Il nuovo settore delle imprese biotecnologiche. Fonti informative, indicatori statistici e ambiti di policy', *L'industria*, vol. 29, no. 2, pagg. 329-346.

d'Amore, R & Vittoria,, P 2009, 'Assessing statistical standards for emerging industries. Applying OECD statistical codes to Italian biotech population lists', *World Review of Science, Technology and Sustainable Development*, vol. 6, no. 2/3, pp. 233-243.

Dasgupta, P & David, P 1994, 'Toward a New Economics of Science', *Research Policy*, vol. 23, no. 5, pp. 487-521.

D'Este, P, & Iammarino, S 2010, The Spatial Profile of University-Business Partnerships, *Papers in Regional Sciences,* vol. 89, no. 2, pp.335-350.

D'Este, P, & Patel P 2007, University-Industry Linkages in the UK: What Are the Factors Underlying the Variety of Interactions with Industry, *Research Policy*, vol. *36*, no. 9, pp. 1295-1313.

DeSolla Price, D 1984, The Science/Technology Relationship, the Craft of Experimental Science, and Policy for the Improvement of High Technology Innovation, *Research Policy* vol. 13, no. 1, pp. 3-20.

Frenken, K & Van Oort F 2004, The geography of research collaboration: theoretical considerations and stylised facts in biotechnology in Europe and the Unites States, in COOKE P. and PICCALUGA A. (Eds.) *Regional Economies as Knowledge Laboratories*, pp.38-57, Edward Elgar, Cheltenham (UK).

Gittelman, M 2005, What makes research socially useful? Complementarities between in-house research and firm-university collaboration in biotechnology, *Revue d'Economie Industrielle*, vol. 110, no. 1, pp. 57-73.

Gittelman, M & Kogut, B 2003, 'Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citations Patterns', *Management Science,* vol. 49, no. 4, 366-382.

Iorio, R, Labory, S & Paci, D 2012, 'The Research Collaboration and Publication's Citation Impact: Evidence in the Italian Biotech Sector', *Rivista Italiana degli Economisti*, vol. 18, no. 1, pp. 75-100.

Katz, J 1994, 'Geographical Proximity and Scientific Collaboration', *Scientometrics*, vol. 31, no. 1, pp. 31-43.

Krackhardt, D & Stern, N 1988, 'Informal Networks and Organizational Crisis: An Experimental Simulation', *Social Psychology Quarterly*, vol. 51, no. 2, pp. 123-140.

Lewison, G & Paraje, G 2004, 'The Classification of Biomedical Journals by Research Levels', *Scientometrics*, vol. 60, no. 2, pp. 145-157.

Liang, L & Zhu, L 2002, 'Major Factors Affecting China's Inter-Regional Research Collaboration: Regional Scientific Productivity and Geographical Proximity', *Scientometrics*, vol. 55, no. 2, pp. 287-316.

McKelvey, M, Alm H &, Riccaboni, M 2003, 'Does Co-location Matter for Formal Knowledge Collaboration in the Swedish Biotechnological-Pharmaceutical Sector?', *Research Policy*, vol. 32, no. 3, pp. 483-501.

Murray, F & Stern, S 2007, 'Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis', *Journal of Economic Behaviour and Organization*, vol. 63, no. 4, pp. 648-687.

Muscio, A 2012, University-Industry Linkages: What Are the Determinants of Distance in Collaborations?, *Papers in Regional Sciences*, DOI: 10.1111/j.1435-5957.2012.00442.x.

Noteboom, B, Van Haverbeke, W, Duysters, G, Gilsing, V & Van den Oord, A 2012, Optimal Cognitive Distance and Absorptive Capacity, *Research Policy, vol.* 36, pp. 1016-1034.

OECD (1996). The Knowledge-Based Economy, Paris.

Pavitt, K 1984, Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory, *Research Policy* vol. 13, no. 6, pp. 343-373.

Ponds, R, Van Oort, F, & Frenken, K 2007, The Geographical and Institutional Proximity of Research Collaboration , *Papers in Regional Science*, vol.86, no.3, pp. 423-443.

Powell, W, Koput, K, & Smith-Doerr, L 1996, Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology*, Administrative Science Quarterly* vol. 41, no. 1, pp.116-145.

Scherngell, T, & Barber, M 2009, Spatial Interaction Modelling of Cross-regions R&D Collaborations: Empirical Evidence from the 5th EU Framework Programme, *Papers in Regional Science,* vol. 88, no. 3, pp.531-546.

Scherngell, T & Barber, M 2011, Distinct Spatial Characteristics of Industrial and Public Research Collaborations: Evidence from the Fifth EU Framework Programme, *Annals of Regional Science*, vol. 46, no.2, pp.247-266.