DiSES Working Papers

Università degli Studi di Salerno
Dipartimento di Scienze Economiche e Statistiche
>>>**www.dises.unisa.it**

New insights on the relationship between geographic and institutional distance in research collaborations: a long period analysis

Rosamaria D'Amore
Roberto Iorio

Working Paper No. 3.229

ISSN: 1971-3029

Dipartimento di Scienze Economiche e Statistiche
Università Degli Studi di Salerno
Via Ponte Don Melillo – 84084; Fisciano (SA) – Italy

Tel +39-089-96.21.55
Fax +39-089-96.20.49
E-mail dises@unisa.it
Web www.dises.unisa.it

NEW INSIGHTS ON THE RELATIONSHIPS BETWEEN GEOGRAPHIC AND INSTITUTIONAL DISTANCE IN RESEARCH COLLABORATIONS: A LONG PERIOD ANALYSIS

Rosamaria D'Amore[1] Roberto Iorio[2]

Abstract

This paper analyses the relationship between institutional and geographic distance in scientific collaborations, evaluating the possible changes when a long period (sixteen years) is taken into consideration and discussing the use of some alternative measures of institutional distance. The main result, obtained by analysing the publications of the Italian biotech firms, is that international publications present an 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

Acknowledgements. The authors are grateful to Pietro Panzarasa for his precious suggestions and encouragement. This paper derives from a strict collaboration among 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; section 3 and 5.2 by Roberto Iorio. Introduction, section 5.3 and Conclusions have been jointly written.

[1] DISES, Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy, rmdamore@unisa.it
[2] DISES, Via Giovanni Paolo II, 132, 84084, Fisciano (SA), Italy, riorio@unisa.it

1. Introduction

It is widely acknowledged that, in the modern knowledge-based economy, therefore, particularly in science-based sectors, the innovation usually derives from collaboration of different agents, often located in different regions or countries and often belonging to different institutions (OECD, 1996). The patterns of collaboration across regions or countries, the ease given by modern information technology to distant collaborations and the persistent importance of vis-à-vis collaboration are themes widely investigated in the literature (KATZ, 1994; LIANG and ZHU, 2002; MCKELVEY et al., 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); anyway collaboration among different innovative institutions happens in several different ways (D'ESTE and PATEL, 2009) and it may be fruitful in terms of quality, because of the existence of complementarities (BONACCORSI and THOMA, 2007; IORIO et al., 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.

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 is 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 measures.

The phenomenon of research collaboration is observed through the lens of co-autorship 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-organizational collaborations. As POWELL et al. (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, the expertise useful for the firm: it is possible to find it in the universities, in the research centres, in the hospitals. The new knowledge generated by these collaborations not only takes the form of industrial innovations, but it is often disclosed trough 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 the authors of the publications belong to are classified in four categories (firms, universities, hospitals and research centres) and their localization is registered too, such as it is possible to identify the geographical extension of the collaborations, distinguishing among regional, extra-regional and international publications. As this research deals with the institutional distance, it is needed a measure of it: according to the existing literature (see, for instance, PONDS et al., 2007), a collaboration between two institutions of the same kind (e.g. two firms or two universities) is characterised by 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 measures based on the

assumption that an increase in the variety of institutions involved in a collaboration increases the institutional distance. Once established the measures of institutional distance, we 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 also possible to control for other factors that may have an effect on the institutional 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 to observe the evolution overtime of institutional and geographic distance, of their relationship and of other significant effects.

The same database used for this analysis has been analysed in a previous study (D'AMORE et al., 2013), which has the same basic research question too (what kind of relationship exists between institutional and geographic distance?). The contribution provided by the present paper consists in the analysis of a longer period of time (the paper by D'AMORE et al., 2013, analyses only the publications from 2003-2005), so that it is also possible to observe the evolution of phenomena in a quite long period of time, and in the use and comparison of several measures of institutional distance. Another novelty of this paper consists in the formulation of two alternative hypotheses regarding the relation between institutional and geographic distance; the first one, supported by the existing literature, is based on a "resource effect": as both kind of distance imply a cost, a trade-off emerges; the alternative hypothesis is based on a "competence effect": 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 and geographic distance move in the same direction. These effects may of course co-exist and may have different strength at different geographic levels, The results of the empirical analysis are of course interpreted in the light of these hypotheses.

The paper is structured in the following way: the second section presents a review of the more relevant literature on the theme of institutional and geographic distance and their relationship in scientific collaborations; the third section presents some hypotheses that may be formulated regarding the relation 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 and geographic distance may be found in PAVITT (1984) and DESOLLA PRICE (1984), who assumed that collaboration between academic and non-academic organizations was more localised into space than collaboration between universities. BOSCHMA (2005) explicitly states 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 long geographic distance. An important reference point for the present study is the paper by PONDS et al. (2008), as they analyse the role of geographical proximity for scientific research collaboration in science-based technologies between three kinds of institutions: academic organizations, firms and governmental/non-profit organizations. They observe the co-autorships in 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 of different kind (institutional distance) Their conclusion is that the collaboration involving different kinds of institutions is more localised than collaboration between the same kind of institutions: when institutional distance increases, spatial distance reduces.

The paper of MCKELVEY et al. (2003) is of great importance for the present paper too, as the content of their study is similar as before and the sector they analyse is the same of this study (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 (firms with universities) than for collaboration among the same kind of institutions (firms with firms; 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 Italian biotech sector; they identify four kinds of institutions (firms, universities, research centres, and hospitals) the authors of the papers belong to and provide a measure of institutional distance, adapting to this purpose the E-I index, a measure 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 randomizations of the institutions. They find results consistent with the trade-off hypothesis, as the publications deriving from international collaborations are 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, 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 if, 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 to establish 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, as, 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 physical sciences in UK, find that, in university-industry partnerships, an high quality of the university departments increase the probability to attract distant business partners; but this holds until a certain threshold of research excellence: beyond such threshold, collaboration with industry tend 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 to collaborate in R&D projects. The analysis is conducted among participants to 290 research projects submitted for the 6th EU Framework Program in 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 to collaborate, 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 organizations) 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 to collaborate between different organizations, but other factors may act in the same way; such factors are: 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, that is reviewed in the previous section, suggests some hypotheses: following PONDS et al. (2008) it is possible to argument 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 is the spatial distance, the less is 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: 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 effects 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 difficultly substitutable and they must be taken "wherever they are".

It is of course possible that the two effects coexist; this may generate ambiguous results, that is a not clear direction of the relationship between institutional and geographic distance. It is also possible 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); in such a case, direction and strength of the relationship may vary with the geographic scale.

4. Data

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

RP Biotech data base. It is a collection of potentially all the Italian firms belonging to the biotech sector, active at 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 ISI-Web of knowledge reports, among other kinds of information, the name of the authors and the name of the institutions the authors belong to. 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 which category each institution belongs to.[1]

Besides, the papers were classified according the nature and the stage of the research, if basic or applied, according to the criterion suggested by LEWISON and PARAJE (2004).

5. Empirical results

5.1 Measures of institutional distance and bivariate analysis

In the empirical analysis two measures are extensively used:

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

An example may be useful to clarify their meaning: if a paper has been written by authors belonging to: University of Milan, University of Rome, Hospital of Turin, firm Rossi S.P.A and 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 of 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 in three categories: national regional publications (written by authors all belonging to institutions located in the same Italian region), national extra-regional papers (all the authors belong to Italian institutions, but at least two institutions are located in two different Italian regions) and international papers (at least one author belongs to a non-Italian institution).

In the following analysis publications written not in collaboration (written by authors belonging only to an Italian biotech firms) were excluded, therefore the analysis is done on 3205 publications. Among these, 620 (19.3%) are national publications, 1399 (43.6%) are extra-regional publications, 1186 (37%) are international publications.[2] Table 2 reports the geographical nature of the publications in the whole periods 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 measure 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 institutions, but, except for the paper by D'Amore et al. (2013), no measures 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-autorship between (authors belonging to) one university, one hospital and one firm implies more institutional distance than a co-autorship between (authors belonging to) two universities and one firm. Therefore KINDINST, that measures exactly how many kinds of institutions are involved in the co-autorship, may be a good measure of institutional distance. As four kinds of institutions were identified, this index ranges from 1 to 4.

KINDINST is an absolute measure 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 an higher probability to include an higher number of institutions (an international paper may include national and foreign institutions; a national paper may include only national institutions), international papers will have an 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 an 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 dimension of publications are in line with the expected results, as the mean values increases 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 needed a measure of institutional distance that "controls" for the increase in the number of institutions. In D'AMORE et al. (2013) this problem is overcome comparing the observed institutional variety with the expected value generated by random combinations of the institutions. In this paper some different measures of institutional distance are provided that control for the number of institutions, 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 institutions, to four, etc (when NUMINST is equal to three, four, etc.).

Table 5 reports the values of KINDINST on the rows for regional, extra-regional, 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 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 a certain evidence between extra-regional and international scale when NUMINST is equal to two and three (the value of KINDINST decreases respectively 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 institutions number across 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 measure of institutional distance derives, in its definition and terminology, from the social network analysis (KRACKARDT and STERN, 1998; D'AMORE ET AL., 2013): we considered the institutions involved in the publications as "nodes" of a network; if, for instance (the 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 kind (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 measure 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 institutions (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, the fourth row the mean value for all the papers; the first column reports the values for all the period, 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 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 calculating 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 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 to 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 to 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/geographic trade off. On the other side, the expected value of the *E index* is not independent on the number of involved institutions, therefore its value, at different geographic scales, may depend not only by the institutional variety (that is what it is desired to measure), but also by 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, having a measure of institutional distance as dependent variable, a measure of geographic distance as independent variable and introducing, as control variables, as many dummy variables as each value of NUMINST: in such a way it is imposed a restriction (the relationship between geographic and institutional distance is the same whatever the number of involved institutions), but it may be obtained a synthetic measure of the relationship, controlling for the number of institutions.

In addition to controlling for the number of institutions, the regression analysis lets to take into account other factors that may influence the relationship between institutional 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, therefore 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 according to the criterion suggested by LEWISON and PARAJE (2004). 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 time dummy variable may let 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 institutional and 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 unit of analysis is the single publication;

-the dependent variable is a measure of institutional distance (KINDINST, *E index* or *E index for firms*);

-the determinant independent variable 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 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 whom the authors belong to, 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, 0 otherwise (the variable NUMINST_3 has value 1 if the paper has been written by authors belonging to three institutions, 0 otherwise; NUMINST_4 has value 1 if the paper has been written by authors belonging to four institutions, 0 otherwise; NUMINST_5 has value 1 if the paper has been written by authors belonging to five institutions, 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 if the paper is of basic nature or not (BASIC assumes value 1 if the paper is of basic nature, 0 otherwise, that is if 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[5] (*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 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 and geographic distance if international and national level are compared, while there is no statistical robust confirm to the hypothesis of the trade-off if regional and extra-regional papers are compared.[6] 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 measure 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 effect), which may have different strength 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 the need to save resources on the institutional side: the resource based effect, that 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.

As easily predictable, the signs of 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/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 an 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 on the other coefficients: the magnitude of coefficients of 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). 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 and geographic distance is confirmed if national and international level 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 an higher institutional variety than early years of the considered period. (tab.8)

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 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, that 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 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 level 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 measure of institutional distance and controlling for the number of institutions, it is possible to find a trade-off between institutional 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 measures of institutional distance, the trade-off between institutional and geographic distance is found whatever the geographic level compared;

- in a multivariate analysis, using KINDINST, *E index* and *E index for firms* as measures 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:

a) it is possible to find a trade-off between institutional and geographic distance if international and national papers are compared;

b) the trade-off between institutional and geographic distance was more intense in the early period (first five years) than in following years.

c) basic research publications are characterised by a greater institutional distance than applied and mixed (basic/applied) ones;

d) 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);

e) 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 analyzed, in order to understand the relation existing between geographic and institutional distance.

The existing literature, including one previous study on the same database but for a more limited time period and with different measures, 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 among 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 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 not an universally accepted definition and measure, several ways to give a measure of it are suggested and such different measures are used to empirically investigate our research question.

Using such different measures of institutional distance, both bivariate and multivariate analysis are conducted, controlling for several factors that may influence the institutional distance: nature of the research, time trends, firm specificity. The empirical analysis confirms the resource based theory, as it is found an inverse relationship between institutional 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 geographic levels (regional and extra-regional) the direction of the relationship is not statistically clear, suggesting the possible contemporary presence of an opposite effect (the competence based effect). If the presence of these two effects is true, the found reducing strength overtime of the trade-off could be explained by an increasing importance of the second effect.

Another interesting result concerns the nature of the research and its increasing effect on the institutional variety: in recent years publication with a basic research content showed an higher institutional variety than more applied papers.

It must be added that this empirical analysis has some limits, as there is not a fine measure of geographic distance and several important variables are not taken into consideration (source of funds, with possible different incentives to institutional collaboration, dimensions of the institutions, etc.). 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 measures and techniques, show that an initial intuition about the relationship between different kinds of distance is becoming a consolidated point. Further and more in depth analyses, that keeps other elements into consideration, need now to be conducted on this topic, as it may shed more light on the way knowledge flows in an innovative sector and 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 measures in a field that is so important for the competitiveness in the contemporary knowledge-based economy.

Notes

[1] A more detailed description of the biotech sector, of the data and more statistical information may be found 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, not only physical distance among the authors and the institutions increases, but differences in norms, culture and languages increase too; all such differences represent barriers to overcome with a "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 between two towns located near the national borders, but one in Italy one 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 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 and little minorities do not have Italian as their main language, while, except a little region in Switzerland, in no other country Italian is spoken.

[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 and institutional distance, 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 Program 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 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 now missing and should be collected for the further stages of our analysis.

Table 8. Results of the ordered probit-determinants of KINDINST (year 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%

References

C. Autant-Bernard, P. Billand, D. Frachisse and N. Massard (2007). Social Distance versus Spatial Distance in R&D Cooperation: Empirical Evidence from European Collaboration Choices in Micro and Nanotechnologies, *Papers in Regional Science 86,* 495-520.

A. Bonaccorsi and G. Thoma (2007). Institutional Complementarity and Inventive Performance in Nanotechnology, *Research Policy 36*, 251-283.

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

A. Broström (2010). Working with Distant Researchers - distance and Content in University-Industry Interaction, *Research Policy 29*, 1311-1320.

R. D'Amore, R.Iorio, S. Labory and A. Stawinoga (2013). Research Collaboration Networks in Biotechnology: Exploring the Trade-Off Between Institutional and Geographic Distances, *Industry and Innovation 20*, 261-276.

R. D'Amore, R.Iorio and A. Stawinoga (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, 2*, 81-100.

P. Dasgupta and P. David (1994). Toward a New Economics of Science, *Research Policy 23*, 487-521.

P. D'Este P. and S. Iammarino (2010). The Spatial Profile of University-Business Partnerships, *Papers in Regional Sciences 86,* 423-443.

P. D'Este and P. Patel (2007). University-Industry Linkages in the UK: What Are the Factors Underlying the Variety of Interactions with Industry, *Research Policy 36*, 1295-1313.

D. DeSolla Price (1984). The Science/Technology Relationship, the Craft of Experimental Science, and Policy for the Improvement of High Technology Innovation, *Research Policy 13*, 3-20.

K. Frenken and F. Van Oort (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).

M. Gittelman (2005). What makes research socially useful? Complementarities between in-house research and firm-university collaboration in biotechnology, *Revue d'Economie Industrielle 90,* 57-73.

M. Gittelman and B. Kogut (2003). Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citations Patterns, Management Science 49, 366-382.

R. Iorio, S. Labory and D. Paci (2012). Research Collaboration and Publication's Citation Impact: Evidence in the Italian Biotech Sector, *Rivista Italiana degli Economisti 18*, 75-100.

J. Katz (1994). Geographical Proximity and Scientific Collaboration, *Scientometrics 31*, 31-43.

D. Krackhardt, N. Stern (1988). Informal networks and organizational crisis: an experimental simulation, *Social Psychology Quarterly, 51*, 123-140.

G. Lewison, G. Paraje (2004). The Classification of Biomedical Journals by Research Levels, *Scientometrics 60*, 145-157.

L. Liang, L. Zhu (2002). Major Factors Affecting China's Inter-Regional Research Collaboration: Regional Scientific Productivity and Geographical Proximity, *Scientometrics 65*, 287-316.

M. McKelvey, H. Alm and M. Riccaboni (2003). Does Co-location Matter for Formal Knowledge Collaboration in the Swedish Biotechnological-Pharmaceutical Sector?, *Research Policy 32*, 483-501.

F. Murray, S. Stern (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 63*, 648-687.

A. Muscio (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.

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

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

K. Pavitt (1984). Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory, *Research Policy 13*, 343-373.

R Ponds, F.Van Oort and K. Frenken (2007). The Geographical and Institutional Proximity of Scientific Collaboration Networks, *Papers in Regional Science 86* , 423-443.

W.Powell, K. Koput and L. Smith-Doerr (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, *Administrative Science Quarterly 41*, 116-145.

T. Scherngell and M. Barber (2009). Spatial Interaction Modelling of Cross-regions R&D Collaborations: Empirical Evidence from the 5th EU Framework Programme, *Papers in Regional Science 88*, 531-546.

T. Scherngell and M. Barber (2011). Distinct Spatial Characteristics of Industrial and Public Research Collaborations: Evidence from the Fifth EU Framework Programme, *Annals of Regional Science 46*, 247-266.