



The geographic proximity effect on domestic cross-sector vis-à-vis intra-sector research collaborations

Giovanni Abramo¹ · Francesca Apponi² · Ciriaco Andrea D'Angelo^{1,2}

Received: 2 November 2021 / Accepted: 9 May 2022 / Published online: 21 May 2022
© The Author(s) 2022

Abstract

Geographic proximity is acknowledged to be a key factor in research collaborations. Specifically, it can work as a possible substitute for institutional proximity. The present study investigates the relevance of the “proximity” effect for different types of national research collaborations. We apply a bibliometric approach based on the Italian 2010–2017 scientific production indexed in the Web of Science. On such dataset, we apply statistical tools for analyzing if and to what extent geographical distance between co-authors in the byline of a publication varies across collaboration types, scientific disciplines, and along time. Results can inform policies aimed at effectively stimulating cross-sector collaborations, and also bear direct practical implications for research performance assessments.

Keywords Research collaborations · Geographical and institutional proximity · Co-authorship · Bibliometrics · Italy

Introduction

Policies in developed countries are aimed at improving efficiency both in knowledge creation and knowledge diffusion among actors involved in different stages of the knowledge value chain. The “triple helix” model (Etzkowitz & Leydesdorff, 1998) envisages the close interaction among the public research area, industrial system, and government institutions as the best way to enhance innovation and development of nations. In this regard, university-industry collaboration allows the match between knowledge, sources, and competence of two realities to achieve higher impact in the long term (Briggs, 2015; Briggs & Wade, 2014; Su et al., 2015). Public–private research collaboration is also one of the main

✉ Giovanni Abramo
giovanni.abramo@iasi.cnr.it

Francesca Apponi
francesca.apponi@uniroma2.it

Ciriaco Andrea D'Angelo
dangelo@dii.uniroma2.it

¹ Laboratory for Studies in Research Evaluation, Institute for System Analysis and Computer Science (IASI-CNR), National Research Council of Italy, Rome, Italy

² Department of Engineering and Management, University of Rome “Tor Vergata”, Rome, Italy

channels which favors knowledge transfer, because it achieves both knowledge creation and transfer at once (D'Este & Patel, 2007).

Among various elements that have an influence on the phenomena of knowledge transfer and research collaboration, proximity is generally acknowledged to be a key factor (Boschma, 2005). Proximity is described as “being close to something measured on a certain dimension” (Knoben & Oerlemans, 2006). It facilitates coordination, enables communication among actors and reduces uncertainty. Specifically, the probability of collaborations is shaped by five types of proximities: organizational, institutional, geographical, social, and cognitive (Boschma, 2005). The importance of proximity dimensions varies across different types of interaction (Alpaydin & Fitjar, 2021), and some forms of proximity might compensate the effect of others.

For instance, geographic proximity can work as a possible substitute for institutional proximity (Crescenzi et al., 2017). In cross-sector collaborations, university and industry have to face the institutional differences that could influence their interaction. The closeness among partners might facilitate cross-sector collaboration, with the geographic proximity compensating for the absence of other proximities, since it increases the possibility of personal interaction and the transfer of tacit knowledge.

The purpose of this work is to investigate the geographic proximity effect on cross-sector collaborations and contrast it with intra-sector collaborations taking place among researchers belonging to the same sector. Previous studies on the geographic distance between inter-institutional collaborations have investigated either type of collaborations alone, but never at once, that is observing the same actors, in the same environment and time period.

In this study, we compare the relevance of the proximity effect for three types of collaboration: public-public, public-private and private-private. We further distinguish between “national only” and “international also” collaborations, to understand if the presence of an international partner might influence the average distance between domestic partners. We also distinguish among the different scientific disciplines, since the intensity of public-private collaboration varies across research fields (Abramo et al., 2021). Finally, we investigate whether the geographic proximity effect varies along time, as it occurs in the case of knowledge spillovers, where it decays over time (Abramo et al., 2020a).

In particular, we intend to answer the following research questions:

- How does the proximity effect impact public-private research collaborations?
- Are there any differences with respect to intra sector public-public or private-private collaborations?
- Does the proximity effect vary in the presence of international collaborations, across fields, and along time?

In order to answer these questions, we conduct a large-scale analysis on the Italian 2010–2017 scientific production indexed in Web of Science (WoS). We measure the average distance of all pairs of authors in the by-line of over 335,000 Italian publications, and apply some statistical tools for analyzing the relationship between distance and type of collaboration across scientific disciplines. The reason why the analysis is restricted to Italy only, is that reconciling and disambiguating private sector affiliations as distinct from public ones in foreign countries is a formidable task for non-nationals.

Results could inform policies aimed at stimulating cross-sector interaction. They also bear direct practical implications for research performance assessments along the cross-sector collaboration dimension, which might need to account for the geographic proximity

effect, not to disfavor relatively remote institutions (Abramo et al., 2012). We warn the reader that these kinds of studies are, by nature, inevitably domestic in scope, as the geography of the country and localization of organizations therein heavily affect results. Consequently, we recommend caution in generalizing results, or even comparing them with those of other national contexts.

We exploit the past years' inroad of bibliometrics that makes it now possible to expand the scope and period of observation of investigations on the topic. Specific bibliometric methodologies were developed specifically for this purpose (Abramo et al., 2010). However, the advantages come together with an observation bias that the reader should be aware of. As this methodology is based on research publication output, it captures only successful collaborations (otherwise the work would not be published). Moreover, not all co-authored publications reveal a real collaboration, and not all successful collaboration necessarily lead to publications.

In the next section, we review the literature on the influence of the proximity effect on private–public research collaboration. In “[Data and methods](#)” section, we present the methodology and data. In “[Results](#)” section, we show the results of the analysis and, in the last section, we conclude the study with our consideration.

Literature review

A large body of the literature has investigated the particular characteristics that could influence the effectiveness of cross-sector research collaboration: size, sector, and R&D intensity of firms; size and scientific specialization of universities (Spithoven et al., 2019).

A number of investigations on various countries, scope, methods and indicators on the spatial distance between public and private organizations engaging in research collaboration have been conducted (Autant-Bernard et al., 2012; Giuliani & Arza, 2009; Giunta et al., 2016; Hewitt-Dundas & Roper, 2011; Tijssen et al., 2011). A trend noticed by many is that the average distance between partners becomes wider over time (Abramo et al., 2020b; Alpaydın, 2019; Waltman et al., 2011), which reflects the globalization of research.

The influence of geographic proximity on the effectiveness of knowledge transfer, as well as a catalyzer of research collaboration, has been demonstrated in cases when the tacit component of knowledge to be shared is conspicuous. When knowledge is transferable mainly through demonstration and observation, requiring face-to-face interaction among partners, knowledge transfer is more easily achieved if the actors are co-located (Gertler, 2003; Morgan, 2004; Singh, 2005), or when the geographic proximity between partners allots frequent interactions (Fitjar & Gjelsvik, 2018; Garcia et al., 2015; Hong & Su, 2013; Petruzzelli & Murgia, 2020).

Universities tend to collaborate with industries located within a limited geographical distance because of lower coordination costs, higher effectiveness in face-to-face interactions, and based on the existence of a common context (Fitjar & Gjelsvik, 2018). However, it appears easier to overcome geographic than cognitive distance (Arant et al., 2019). In fact, when partners are cognitively close, they tend to interact at larger geographical distances (Garcia et al., 2018).

Moreover, the interplay of geographical distance and quality of the university partners also influences both collaborations and outcomes. Both geographical proximity and research quality appear positively associated with the frequency of university–industry partnerships; however, differences occur across scientific disciplines (D’Este & Iammarino,

2010). The geographic proximity of university and industry might favor research collaboration even if there is a trade-off between the quality of local universities and the higher costs associated with greater distance (Guerrero, 2020; Tang et al., 2020). In the UK, physical co-location with top-tier universities favors cross-sector collaboration. However, if faced with this choice, UK firms (especially the R&D-intensive ones), appear to prefer quality over distance (Laursen et al., 2011). It was shown an inverted-U shape relationship between excellence of university partners and distance with industry partners (D’Este & Iammarino, 2010). In a subsequent study, the same authors found that firms located in intensive R&D clusters tend to partner with universities regardless of their location, while firms outside such clusters tend to partner with local universities (D’Este et al., 2013).

On the contrary, Abramo et al. (2011) revealed a problem of information asymmetry in the market for university-industry research collaboration in Italy. The authors found that, in 93% of cases, firms could have collaborated with a higher quality university. In 54% of cases, there was at least one university both closer and of higher quality when compared to the university that was actually chosen for collaboration. At single professor level, in 95% of cases the private company could have partnered with a higher performing professor in the same field of the collaboration; in 65% of cases, the choice could have been a better-performing professor, located closer to the company. Tijssen et al. (2020) identified a number of determinants affecting university-industry research collaborations, varying across distance zones. Four of them appear to be common to all zones, namely intensive R&D firms, research size of a university and its quality, and gatekeepers among the faculty.

Data and methods

In order to answer the research questions, an econometric analysis of the Italian scientific production of the period 2010–2017 was carried out. The data source is the Italian National Citation Report, extracted from the Web of Science *core collection* imposing “Italy” as affiliation country of at least one author. The unit of observation is the single scientific publication resulting from national extramural collaboration. In WoS, each bibliometric address is composed of two parts: the first one refers to the affiliation and is made up of four “segments”, corresponding in general to the macro-organization (Seg1) and to its internal articulations at the level of “School” (Seg2), “Department” (Seg3) and “Research unit” (Seg4). The second part consists of toponymic information: City, Province, State, Zip_Code and Country. Therefore, in order for a publication to be defined as the result of national extramural collaboration, the following two conditions referring to the byline must be met:

- It must contain at least two authors;
- It must contain at least two Italy addresses referring to distinct organizations, i.e., distinct “Seg1—City” pairs.

We measure the geographic proximity between co-authors of a publication in terms of the average geodesic distance of the cities associated with all distinct Seg1+City pairs found in the publication’s address list.¹ This distance is a function of geographical

¹ Authors with multiple affiliations are counted multiple times. This would entail a distortion in the average value of the distance computed for publications by such authors. However, in Italy multiple affiliations hardly concern national institutions, rather, a national institution and a foreign one, and we do not observe the latter.

Table 1 The address list of a publication in the dataset

Seg1	Seg2	Seg3	Seg4	City	Province	STATE	Zip_code	Country
Univ Pavia	Dept Appl Hlth Sci	Sect Med Stat and Epide- miol		Pavia			I-27100	IT
Univ Pavia	Dept Math			Pavia			I-27100	IT
Ca Granda Osp Maggiore Policlinico	Serv Biostat			Milan				IT
Osped Niguarda	Serv Biostat			Milan				IT

coordinates of cities extracted from the Italian institute of statistics (ISTAT)² for Italian LAUs.³ For reasons of computational complexity, publications with more than ten distinct “Seg1—City” pairs are excluded,⁴ as well as those in which we are unable to geo-locate all the cities indicated in the address list, due to a transcription error in the source data. In total, the analysis dataset includes 335,574 publications. As an example, we report the case of the publication with accession number WOS:000208151600003, whose address list is given in Table 1.

This list consists of three distinct Seg1—City pairs:

- Univ Pavia—Pavia;
- Ca Granda Osp Maggiore Policlinico—Milan;
- Osped Niguarda—Milan.

As these are public research organizations (one university and two hospitals), the publication is classified as “national intra-sector public extramural collaboration”. In case one (or more) public and one (or more) private national organization(s) are recognizable in the address list, the publication is classified as “cross-sector national collaboration”. Disambiguation of public vs private organizations requires manual scrutiny and profound knowledge of the country under observation. A subsequent step is reconciliation of all bibliographic addresses with “Italy” as affiliation country (D’Angelo et al., 2011). Through such reconciliation it is possible to tag a publication as fruit of:

- Intra-sector public collaboration: if all “Seg1—City” pairs related to the Italian addresses, pertain to recognized public national organizations;
- Intra-sector private collaboration: if all “Seg1—City” pairs, pertain to recognized private Italian organizations;
- Cross-sector collaboration: if “Seg1—City” Italian pairs in the address list pertain both to public and to private national organizations.

² <https://www.istat.it/it/archivio/6789>, last accessed on 21 March 2022.

³ The LAU level consists of municipalities or equivalent units. The assumption of the “city” as an element of geo-referencing implies that any collaboration between organizations located in the same city occurs at zero distance.

⁴ These amount to 1.3% of total observations. Even considering that this elimination affects some disciplines more than others, the effect remains very limited in size.

Finally, the possible presence of one or more addresses with a country different from “Italy” implies the international tagging of the publication.

It is understood that the presence/absence of a foreign address does not affect the average value of the distance between the authors of a publication, which is exclusively calculated with reference to the Italian addresses, being the aim of the work the analysis of the geographical proximity in the national extramural collaborations.

In order to deepen the analysis at field level, each publication in the dataset is assigned to the subject category of the hosting journal.⁵ Considering the aggregation of subject categories in macro-areas,⁶ Table 2 shows the breakdown of the 335,574 publications in the dataset by type and macro-area.

Results

We will initially present the descriptive analysis of geographic distances in the different types of collaboration and at the macro-area level. Next, we will illustrate the results of an inferential analysis.

Descriptive analysis

The distribution of the average distance values between the co-authors of publications resulting from different types of collaboration, as shown in Fig. 1, reveals already at a glance the presence of a differentiated proximity effect for cross-sector versus intra-sector collaborations. The former have higher values of both central tendency (mean and median) and interquartile distance. Given the geography of the country, the maximum values (all around just over 1000 km) obviously tend to saturate. Table 3 reports the full descriptive statistics of the average distance for the publications in the dataset. For each analyzed set, the high skewness determines a very significant deviation of the mean values from the medians. In particular, intra-sector public national collaborations show an average distance of 132.7 km and a median distance of 47.1 km. Intra-sector private national collaborations have longer distances, with an average of 147.3 km and a median of 50.3 km. The figure for cross-sector national collaborations rises further to 148.2 km and 80.4 km for their mean and median, respectively. The presence of a foreign partner seems to show a significant effect on the mean/median distance of private intra-sector and cross-sector collaborations, but no effect on public intra-sectors.

It should also be noted that at least a quarter of the total number of publications is the result of collaboration between researchers located in the same city, even if in different organizations, as indicated by the value of the 25th percentile of the distribution of distances, invariably zero for all the sets under analysis, except for cross-sector international collaborations (where the average distance is less than 8 km). Regarding the variability, the distributions of intra-sector public and cross-sector collaborations show very similar values of standard deviation, around 180–190 km. In contrast, private

⁵ In the case of multi-category journals, the publication is assigned indiscriminately to all subject categories (multiple counting).

⁶ Our assignment of SCs to macro-areas follows a pattern previously published on the website of ISI Journal Citation Reports, but no longer available on the current Clarivate portal.

Table 2 Analysis dataset by type of publication and macro-area

	Intra-sector public (national)	Intra-sector public (international)	Cross-sector (national)	Cross-sector (international)	Intra-sector private (national)	Intra-sector private (international)	Total
Mathematics	5724	3296	198	87	2	1	9308
Physics	22,442	32,142	2151	803	41	29	57,608
Chemistry	14,245	10,207	1246	506	20	10	26,234
Earth and Space Science	12,484	10,494	1161	527	16	12	24,694
Biology	31,297	19,894	1849	1002	26	34	54,102
Biomedical Research	37,119	20,669	1414	891	22	24	60,139
Clinical Medicine	69,385	37,044	1678	906	11	22	109,046
Psychology	2567	1729	34	17	0	0	4347
Engineering	37,084	23,191	5977	2078	143	106	68,579
Economics	4381	2708	204	120	9	3	7425
Law, Political and Social Science	4475	2543	157	67	2	2	7246
Art and Humanities	1832	757	106	43	1	0	2739
Multidisciplinary	623	553	25	32	3	2	1238
Total*	189,632	128,438	11,940	5157	217	190	335,574

*The total is less than the sum of column data, due to multiple counting of publications in SCs falling in more than one macro-area

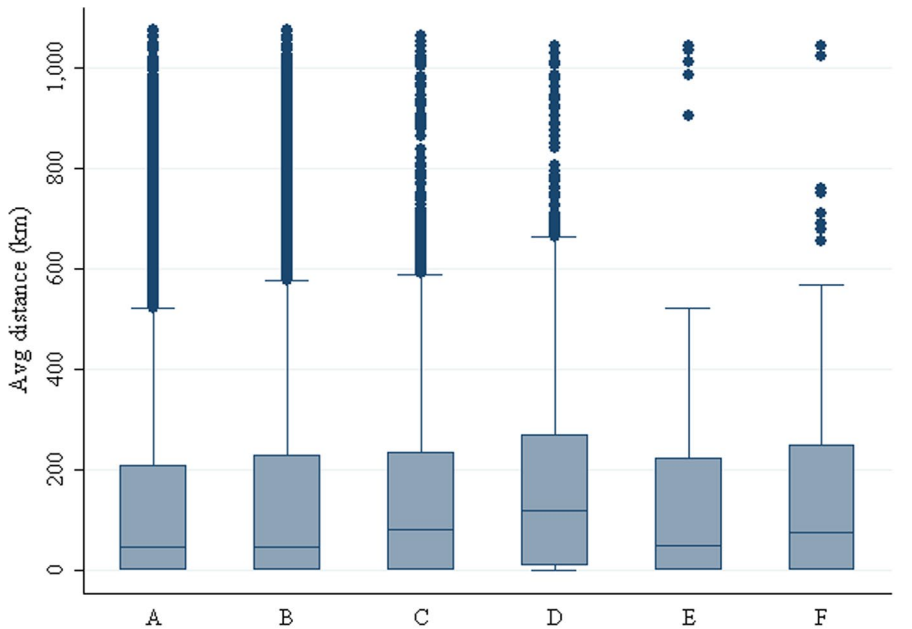


Fig. 1 Box plot of average distance of publications' co-authors, by collaboration type. A=Intra-sector public national; B=Intra-sector public international; C=Cross-sector national; D=Cross-sector international; E=Intra-sector private national; F=Intra-sector private international

Table 3 Descriptive statistics of distances (in km) between publications' co-authors, by collaboration type

	Intra-sector public national	Intra-sector public international	Cross-sector national	Cross-sector international	Intra-sector private national	Intra-sector private international
Obs	189,632	128,438	11,940	5157	217	190
Mean	132.7	136.6	148.2	171.3	147.3	166.1
Std Dev	184.1	183.8	178.4	191.0	210.3	219.8
Max	1074.9	1074.9	1063.8	1042.8	1042.8	1045.0
Skewness	1.780	1.596	1.631	1.504	2.071	1.669
Kurtosis	6.291	5.549	6.058	5.676	7.881	5.844
25%	0	0	0	7.8	0	0
50%	47.1	45.6	80.4	118.8	50.3	74.4
75%	209.2	230.5	236.1	270.1	222.8	249.3
90%	396.0	402.8	392.5	442.7	468.6	477.4
95%	513.6	512.0	502.0	528.0	511.7	567.7

intra-sectors show significantly greater variability, with a standard deviation of 210 km for nationals and 220 km for internationals. Regarding the international dimension of the collaboration, it seems to show a significant effect on the average distance of Italian co-authors of a private intra-sector and cross-sector publication. In contrast, in the absence of a private organization in the byline, the presence of a foreigner does not

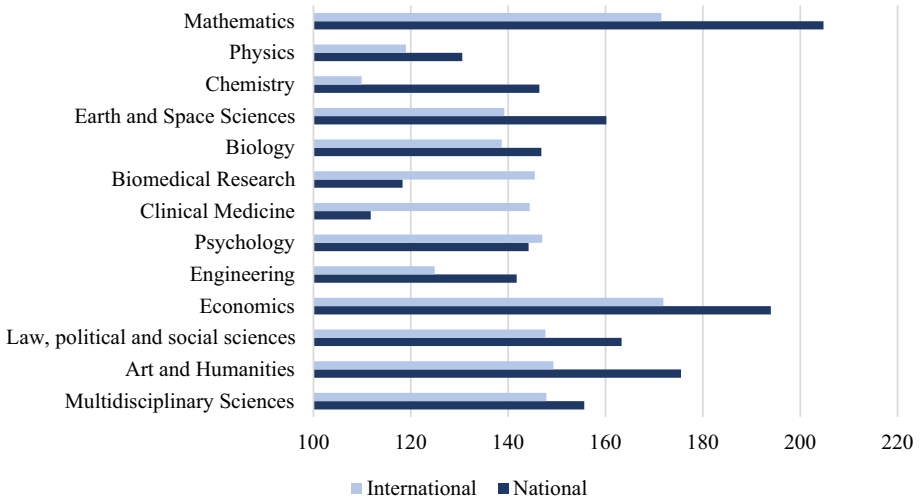


Fig. 2 Average distances (in km) between co-authors of intra-sector public publications by macro-area. (Color figure online)

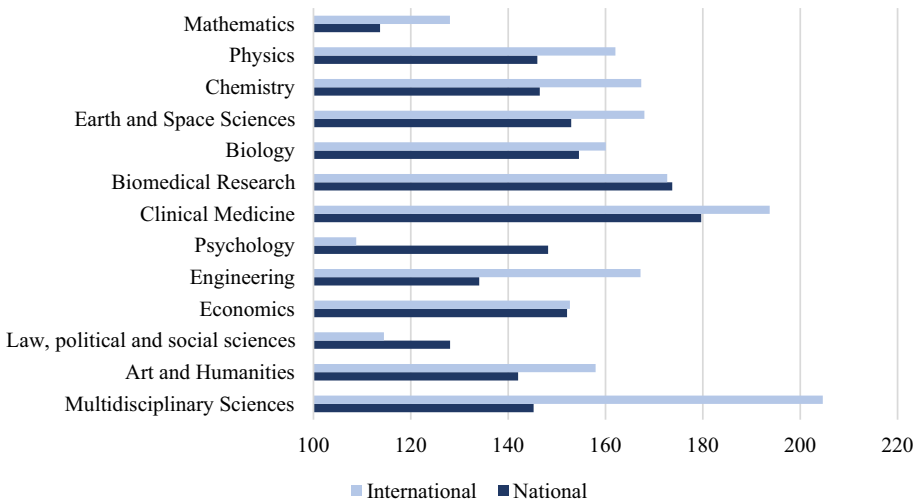


Fig. 3 Average distances (in km) between co-authors of cross-sector publications by macro-area. (Color figure online)

appear to be associated with a change in mean/median distance between the co-authors of a publication.

The following figures show the breakdown of the data by macro-area. Figure 2 shows the average distances for publications that are the result of intra-sector public collaborations, and it can be seen that, for the national case, the greatest distances are in Mathematics, Social Sciences (in particular Economics) and Art and Humanities. The shortest distances are in Clinical Medicine and Biomedical Research. In these two areas, the presence of at least one author with a foreign affiliation significantly increases the average distance between partners. This is also occurring in Psychology, while the opposite is true in all

other areas. Cross-sector collaborations are a different matter: Fig. 3 shows that, for the national case, maximum average distances are found in Clinical Medicine and Biomedical Research, while minimum ones are found in Mathematics and Law, Political and Social Sciences. The presence of authors with foreign affiliation increases the average distance between cross-sector partners in all areas but three: Law, Political and Social Sciences, Psychology, Biomedical Research.

Institutional proximity vs geographic proximity

As argued in the Introduction, geographic proximity can work as a possible substitute (or complement) for institutional proximity. Institutions involved in cross-sector collaborations have to face and manage their differences in organizational cultures, practices, objectives, motivations, incentives, backgrounds. Such differences could influence their interaction and impact on the final outcome of the collaboration itself. The geographic closeness among partners might mitigate the effect of other distances, since it increases the possibility of personal interaction and the transfer of tacit knowledge. The descriptive analyses presented in the previous section indicate that the distributions of the average distances between co-authors of publications resulting from different types of collaboration reveals already at a glance the presence of a differentiated proximity effect for cross-sector versus intra-sector collaborations. Next we test the statistical significance of these differences considering institutional proximity proxied by a dummy variable assuming value 1 for cross-sector collaborations, and 1 for intra-sector ones.

At this purpose, Table 1 shows data related to a two-group mean-comparison test, revealing that the observed difference between the mean distances of the two sets of publications is statistically significant at overall level and in 9 out of 13 total areas. The negative value of the t statistics indicates that the average distance between co-authors in cross-sector collaborations is higher than in intra-sector: this is the case at the overall level and in all STEM areas but Mathematics.

Area	t	Degrees of freedom	Pr ($ T > t $)
Art and Humanities	1.252	2737	0.211
Biology	-3.644***	54,100	0.000
Biomedical Research	-12.136***	60,137	0.000
Chemistry	-4.588***	26,232	0.000
Clinical Medicine	-17.716***	109,044	0.000
Earth and Space Sciences	-1.478	24,692	0.139
Economics	2.886***	7423	0.004
Engineering	-3.076***	68,577	0.002
Law, political and social science	2.584**	7244	0.010
Mathematics	5.548***	9306	0.000
Multidisciplinary Sciences	-1.027	1236	0.305
Physics	-7.880***	57,606	0.000
Psychology	0.408	4345	0.683
Overall	-16.496***	432,703	0.000

In order to provide a definite answer, we perform a logit regression on the dataset. The logit model is specified assuming that distance between potential collaborators affects

Table 4 Average values of the regression model variables

Variable		Mean	Std dev.	Max	Median	Skewness	Kurtosis
Y	Cross-sector (baseline intrasector)	0.051	0.220	1	0	4.084	17.68
X ₁	Avg distance	0	1	5.106	-0.472	1.699	5.982
X ₂	International	0.399	0.490	1	0	0.414	1.17
X ₃	No of authors	0	1	157.5	-0.116	49.392	5750.85
X ₄	Presence of universities	0.846	0.361	1	1	-1.913	4.66
X ₅	Period (baseline 2010–2013)	0.564	0.496	1	1	-0.256	1.07

Table 5 Correlation matrix of the variables of the regression model

	Y	X ₁	X ₂	X ₃	X ₄	X ₅
Y	1					
X ₁	0.025	1				
X ₂	-0.046	0.012	1			
X ₃	-0.008	0.063	0.140	1		
X ₄	-0.008	0.104	-0.100	0.013	1	
X ₅	0.037	0.028	0.044	0.013	0.021	1

Y cross-sector collaborations, X₁ avg distance, X₂ international, X₃ no. of authors, X₄ presence of universities, X₅ period

whether collaboration actually happens. Therefore, we consider a binary response variable assuming value 1 for publications resulting from cross-sector collaborations (and 0 as baseline, for publications resulting from intra-sector collaborations), depending on:

- The average distance between co-authors (X₁).
- Presence of at least one foreign affiliation, specified by a dummy variable (X₂).
- Number of authors in the byline (X₃).
- Presence of at least one university, specified by a dummy variable (X₄).

In order to estimate a possible temporal pattern in the data, we consider an additional dummy (X₅), assuming value 1 for 2014–2017 publications, and 0 for 2010–2013 ones. Furthermore, X₁ and X₃ are expressed applying a z-score transformation. For example, a z-score of 1.2 for X₁ indicates that the average distance of co-authors of the publication is 1.2 standard deviations higher than the average value measured on the whole dataset.

Finally, in order to control for area effects, we also consider 13 additional dummies, one for each macro-area.

Table 4 reports the descriptive statistics of the model variables and Table 5 the correlation indexes between pairs of variables.

The fourth column of Table 4 indicates that publications resulting from cross-sector collaborations represent just 5.1% of the total. Those with at least one foreign author are just under 40%. 84.6% of the publications have at least one address attributable to a national university. As for the average distance (X₁) and the number of authors (X₃), Table 4 shows nihil average values due to the z-score transformation. Finally, we notice a slight imbalance of the dataset in the second period: the publications of the 4-year period 2014–2017 are 56.4% of the total, against 45.6% of 2010–2013.

Table 6 Logit regression; dependent variable: 1 for publications resulting from cross-sector collaborations, 0 otherwise (Model 2 embeds area effects with “Physics” as baseline)

	Model 1			Model 2		
	Coeff	Std Err	z	Coeff	Std Err	z
_cons	0.060***	0.001	−123.06	0.063***	0.002	−103.03
Avg distance	1.116***	0.008	15.08	1.098***	0.007	14.78
International	0.619***	0.011	−27.43	0.607***	0.009	−33.14
No of authors	0.979*	0.012	−1.76	1.018***	0.005	3.56
Presence of universities	0.816***	0.017	−9.52	0.889***	0.016	−6.4
Period	1.447***	0.024	22.53	1.398***	0.020	23.54
Art and Humanities				0.890	0.077	−1.35
Biology				0.934**	0.025	−2.52
Biomedical Research				0.669***	0.019	−14.05
Chemistry				1.234***	0.039	6.71
Clinical Medicine				0.407***	0.011	−32.52
Earth and Space science				1.233***	0.039	6.59
Economics				0.741***	0.045	−4.99
Engineering				2.201***	0.050	34.92
Law, Political and Social Science				0.513***	0.036	−9.44
Mathematics				0.520***	0.033	−10.33
Multidisciplinary				0.814	0.112	−1.5
Psychology				0.201***	0.029	−11.27
Number of obs		335,574		432,705		
LR χ^2		1019.33		9398.97		
Prob > χ^2		0.000		0.000		
Log likelihood		−67,040.9		−85,974.8		
Pseudo r-squared		0.0075		0.0518		

Statistical significance: *p-value < 0.10, **p-value < 0.05, ***p-value < 0.01

Table 5 reveals a practically non-existent correlation between the variables of the model, which leads us to exclude possible multicollinearity effects: the highest coefficient (0.140) concerns the X_2 – X_3 pair, indicating a very weak link between the international character of the publication and the number of its authors.

Table 6 shows the results of the logit regression. The estimated coefficients are expressed in terms of odds ratios: the interpretation of the β_i OR is as follows:

- For a dummy independent variable, the change from the baseline (0) to the reference value (1) is associated with a $\beta_i \times 100\%$ variation in the probability that the publication results from a cross-sector collaboration.
- For a continuous variable, after normalization through z-score, a one-standard deviation increase in the value of the variable is associated with a $\beta_i \times 100\%$ variation in the probability that the publication results from a cross-sector collaboration.

The left side of Table 6 shows data related to the overall analysis (Model 1), while the right side shows those obtained considering area effects (Model 2). The coefficients are all

significant in Model 1, while in Model 2 area effects are not significant for *Art and Humanities*, and *Multidisciplinary*.

Other things being equal and compared to publications resulting from intra-sector collaborations, the probability that a publication results from a cross-sector collaboration increases by 11.6% when the average distance between authors rises by one standard deviation in Model 1. This probability decreases to 9.8% when considering area effects (Model 2). Also, X_5 (Period) presents a positive effect on the probability of a cross-sector research collaboration (+44.7% in Model 1, and +39.8% in Model 2).

All other covariates under examination have negative impacts on the dependent variable Y (institutional proximity) with all odds ratios below 1. In particular, the probability of cross-sector collaboration decreases by 38.1% with the presence of foreign co-authors in Model 1, and by 39.3% in Model 2. Furthermore, the presence of an academic researcher in the byline, other things being equal, decreases by 18.4% the probability to have a cross-sector collaboration with respect to have an intra-sector one. The magnitude of this effect drops to 11.1% in Model 2.

The last column of Table 6 shows that the model betas do not vary significantly when area effects are considered, except for X_2 (number of authors), so all the results and effects highlighted with the specification of Model 1 are repeated with that of Model 2. Contrasting outcomes emerge for the number of authors, since its effect is negative in Model 1, positive when considering area effects. In fact, in Model 2 the probability of having a publication resulting from cross-sector collaboration increases by 1.8% when the number of authors grows by one standard deviation.

As for area effects, it is worth noting that compared to the baseline (Physics) cross-sector collaborations are not influenced in the same way. As expected, *Engineering* presents the highest odds ratio (2.201), confirming that this research area increases the probability of cross-sector collaborations. *Chemistry* (+23.4%) and *Earth & Space Science* (+23.3%) are the other two areas with positive impact. The other areas present negative effects on cross-sector collaboration, that is no-significant, as previously remarked, in *Art and Humanities* and *Multidisciplinary*.

Findings confirm the absence of compensation effects. Geographic proximity does not compensate for institutional distance. Collaborations between public and private researchers involve higher distances between partners.

The very low values of R-squares indicate the importance of considering more control variables, especially other dimensions of proximities which bibliometric metadata can hardly capture.

Discussion and conclusions

Public–private research collaborations are one of the most relevant targets of developed countries policies aiming at improving efficiency both in knowledge creation and knowledge diffusion. Understanding the ways in which they are implemented, the motivations behind for involved partners and the factors that hinder them is key to optimizing such policies. Among elements that have an influence on research collaboration, proximity is generally acknowledged to be a key factor. In cross-sector collaborations, the partners belong to different worlds and have to face institutional differences that could heavily influence their interaction. The cultural, motivational and linguistic “distance” between a researcher working in a university or a public research institution and a colleague working in a private

company probably makes it more necessary to resort to face-to-face interactions for the development of the necessary trust, the set up and the tuning of optimal conditions for the achievement of the aims of the collaboration. In such conditions, geographic proximity can work as a possible substitute for institutional proximity. This is at least what we get from scanning previous literature on the subject.

In this study, we tested the presence of this “compensation” effect between geographical and institutional proximity, referring to the Italian context. Our results indicate that this effect is not detectable; on the contrary, in cross-sector collaborations the average distances between partners are greater than in collaborations involving partners from the same sector, i.e., institutionally more similar. One could hypothesize the presence of an “intermediation” effect of quality of the prospect partner. In particular, from the perspective of private firms, especially the R&D-intensive ones, they could prefer quality over distance. This evidence, which has emerged in several studies related to the UK context, has, however, already been refuted for the Italian case. Abramo et al. (2011) have in fact found the existence of an information asymmetry that would prevent Italian firms, in at least half of the cases, to choose as their partners for possible research collaborations excellent researchers closely located to the company headquarters.

Rather, the result that emerged in the study conducted can find an explanation in the specificity of the context of analysis and, in particular, in the non-homogeneous distribution of private R&D activities on the Italian territory. According to the latest statistical survey just published (ISTAT, 2021), more than 75% of Italian R&D expenditure by private companies is concentrated in five of the twenty regions: apart from Lazio (located in the center of Italy), the remaining four are all Northern regions: Lombardy, Emilia-Romagna, Piedmont and Veneto. The whole of the South covers only just over 9% of national business expenditure. These data attest to an evident greater difficulty for a researcher from a university or public research institution in the South to find a potential industrial partner for a research collaboration of mutual interest, located nearby or within the same region. This has obvious implications for research performance assessments of the so-called “third mission” of universities, an assessment to which in Italy a part of the ordinary funding provided by the Ministry of University and Research is linked.

We observe that the average geographic distance between project team members has an influence on cross-sector research collaborations, which are less concentrated and more geographically distributed than intra-sector. Considering the casual effect in our regression model, as the geographic distance among partners of a research project increases, it increases the probability that partners belong to different sectors.

With regard to area effects, the number of authors shows positive effect on the probability that a publication results from a cross-sector collaboration rather than intra-sectors. *Engineering*, *Chemistry* and *Earth and Space science* have positive impact on cross-sector collaborations, particularly *Engineering* which presents the highest odds ratio. On the contrary, the international dimension and the presence of university researchers among co-authors lead to a decrease in the probability that cross-sector collaborations occur. A relevant trend emerging from our study is that the probability of public–private collaborations tends to grow over time under the same conditions. In line with the findings by D’Este and Iammarino (2010) that research quality appears positively associated with the frequency of university–industry partnerships, the above phenomenon might be partly explained by the general improvement of academic research performance, fostered by the introduction of performance-based research funding systems in Italy (Abramo & D’Angelo, 2021).

Differently from what argued by Alpaydin and Fitjar (2021), that some forms of proximity might compensate the effect of others, and in particular that geographic proximity can

work as a possible substitute for institutional proximity (Crescenzi et al, 2017), our findings show that this is not the case in Italy. Certainly, our study is by nature inevitably domestic in scope, as the geography of the country and the private R&D system features heavily affect results, as already widely argued. In particular, in a follow-up study that we have initiated, we intend to verify if and how much the results obtained depend on the geographical distribution/concentration of R&D activities in Italy. Consequently, we recommend caution in generalizing results, or even comparing them with those of other national contexts.

Finally, we cannot help reminding the intrinsic limitations of the bibliometric approach adopted: (i) observing publication's authorships allows to capture only successful collaborations; (ii) not all co-authored publications reveal a real collaboration, and not all successful collaborations necessarily lead to publications. Nevertheless, the authors believe that these limitations are largely counterbalanced by the power of the approach itself in terms of numerosness, a power found in the high level of significance of the analyses conducted on the phenomenon of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abramo, G., Apponi, F., & D'Angelo, C. A. (2021). Public-private research collaborations: Longitudinal field-level analysis of determinants, frequency and impact. *Journal of Economic Surveys*, *35*(5), 1261–1270. <https://doi.org/10.1111/joes.12425>
- Abramo, G., & D'Angelo, C. A. (2021). The different responses of Italian universities to introduction of performance-based research funding. *Research Evaluation*, *30*(4), 514–528.
- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2020a). The role of geographical proximity in knowledge diffusion, measured by citations to scientific literature. *Journal of Informetrics*, *14*(1), 101010.
- Abramo, G., D'Angelo, C. A., & Di Costa, F. (2020b). Knowledge spillovers: Does the geographic proximity effect decay over time? A discipline-level analysis, accounting for cognitive proximity, with and without self-citations. *Journal of Informetrics*, *14*(4), 101072.
- Abramo, G., D'Angelo, C. A., Di Costa, F., & Solazzi, M. (2011). The role of information asymmetry in the market for university-industry research collaboration. *The Journal of Technology Transfer*, *36*(1), 84–100.
- Abramo, G., D'Angelo, C. A., & Solazzi, M. (2010). Assessing public-private research collaboration: Is it possible to compare university performance? *Scientometrics*, *84*(1), 173–197.
- Abramo, G., D'Angelo, C. A., & Solazzi, M. (2012). A bibliometric tool to assess the regional dimension of university-industry research collaborations. *Scientometrics*, *91*(3), 955–975.
- Alpaydin, U. A. R. (2019). Exploring the spatial reach of co-publication partnerships of multinational enterprises: To what extent does geographical proximity matter? *Regional Science*, *6*(1), 281–298.
- Alpaydin, U. A. R., & Fitjar, R. D. (2021). Proximity across the distant worlds of university-industry collaborations. *Papers in Regional Science*, *100*(3), 689–711.
- Arant, W., Fornahl, D., Grashof, N., Hesse, K., & Söllner, C. (2019). University-industry collaborations—The key to radical innovations? *Review of Regional Research*, *39*(2), 119–141.
- Autant-Bernard, C., Billand, P., & Massard, N. (2012). Innovation and space—From externalities to networks. In C. Karlsson, B. Johansson, & R. Stough (Eds.), *The regional economics of knowledge and talent*. Edward Elgar Publishers.
- Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61–74.

- Briggs, K. (2015). Co-owner relationships conducive to high quality joint patents. *Research Policy*, 44(8), 1566–1573. <https://doi.org/10.1016/j.respol.2015.05.011>
- Briggs, K., & Wade, M. (2014). More is better: Evidence that joint patenting leads to quality innovation. *Applied Economics*, 46(35), 4370–4379.
- Crescenzi, R., Filippetti, A., & Iammarino, S. (2017). Academic inventors: Collaboration and proximity with industry. *The Journal of Technology Transfer*, 42(4), 730–762.
- D'Angelo, C. A., Giuffrida, C., & Abramo, G. (2011). A heuristic approach to author name disambiguation in bibliometrics databases for large-scale research assessments. *Journal of the American Society for Information Science and Technology*, 62(2), 257–269.
- D'Este, P., Guy, F., & Iammarino, S. (2013). Shaping the formation of university–industry research collaborations: What type of proximity does really matter? *Journal of Economic Geography*, 13(4), 537–558.
- D'Este, P., & Iammarino, S. (2010). The spatial profile of university–business research partnerships. *Papers in Regional Science*, 89(2), 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*, 36(9), 1295–1313.
- Etzkowitz, H., & Leydesdorff, L. (1998). The endless transition: A “triple helix” of university–industry–government relations: Introduction. *Minerva*, 36, 203–208.
- Fitjar, R. D., & Gjelsvik, M. (2018). Why do firms collaborate with local universities? *Regional Studies*, 52(11), 1525–1536.
- Garcia, R., Araujo, V., Mascarini, S., Gomes Dos Santos, E., & Costa, A. (2018). Is cognitive proximity a driver of geographical distance of university–industry collaboration? *Area Development and Policy*, 3(3), 349–367.
- Garcia, R., Araujo, V., Mascarini, S., Gomes Santos, E., & Costa, A. (2015). Looking at both sides: How specific characteristics of academic research groups and firms affect the geographical distance of university–industry linkages. *Regional Studies, Regional Science*, 2(1), 518–534.
- Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, 3(1), 75–99. <https://doi.org/10.1093/jeg/3.1.75>
- Giuliani, E., & Arza, V. (2009). What drives the formation of “valuable” university–industry linkages? *Research Policy*, 38, 906–921.
- Giunta, A., Pericoli, F. M., & Pierucci, E. (2016). University–industry collaboration in the biopharmaceuticals: The Italian case. *The Journal of Technology Transfer*, 41(4), 818–840.
- Guerrero, D. F. (2020). Industry–university collaboration in rural and metropolitan regions: What is the role of graduate employment and external non-university knowledge? *Journal of Rural Studies*, 78, 516–530.
- Hewitt-Dundas, N., & Roper, S. (2011). Creating advantage in peripheral regions: The role of publicly funded R&D centres. *Research Policy*, 40(6), 832–841.
- Hong, W., & Su, Y. S. (2013). The effect of institutional proximity in non-local university–industry collaborations: An analysis based on Chinese patent data. *Research Policy*, 42(2), 454–464.
- ISTAT. (2021). *La spesa in ricerca e sviluppo*. ISTAT.
- Knoben, J., & Oerlemans, L. A. (2006). Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8(2), 71–89.
- Laursen, K., Reichstein, T., & Salter, A. (2011). Exploring the effect of geographical proximity and university quality on University–Industry collaboration in the United Kingdom. *Regional Studies*, 45(4), 507–523.
- Morgan, K. (2004). The exaggerated death of geography: Learning, proximity and territorial innovation systems. *Journal of Economic Geography*, 4(1), 3–21.
- Petruzzelli, A. M., & Murgia, G. (2020). University–Industry collaborations and international knowledge spillovers: A joint-patent investigation. *The Journal of Technology Transfer*, 45(4), 958–983.
- Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*, 51(5), 756–770.
- Spithoven, A., Vlegels, J., & Ysebaert, W. (2019). Commercializing academic research: A social network approach exploring the role of regions and distance. *The Journal of Technology Transfer*, 46, 1–36.
- Su, C. Y., Lin, B. W., & Chen, C. J. (2015). Technological knowledge co-creation strategies in the world of open innovation. *Innovation*, 17(4), 485–507.
- Tang, Y., Motohashi, K., Hu, X., & Montoro-Sanchez, A. (2020). University–industry interaction and product innovation performance of Guangdong manufacturing firms: The roles of regional proximity and research quality of universities. *The Journal of Technology Transfer*, 45(2), 578–618.
- Tijssen, R., van de Klippe, W., & Yegros, A. (2020). Localization, regionalization and globalization of university–business research co-operation in the United Kingdom. *Papers in Regional Science*, 99(5), 1215–1236.

- Tijssen, R., Waltman, L., & Van Eck, N. (2011). Globalization of public-private research: Monitoring co-publication trends in worldwide science. In *Proceedings of the Atlanta Conference on Science and Innovation Policy*.
- Waltman, L., Tijssen, R., & Van Eck, N. (2011). Globalisation of science in kilometres. *Journal of Informetrics*, 5, 574–582.