The Journal of Technology Transfer (2022) 47:1524–1548
https://doi.org/10.1007/s10961-021-09881-2

The role of the academic relations of former graduate students in university-firm collaboration

Daniel Gama e Colombo[1] · **Renato de Castro Garcia**[1]

Accepted: 6 September 2021 / Published online: 17 September 2021

Abstract

This paper investigates the contribution of the personal ties of former graduate students to university-firm collaboration. Using the proximity framework and the underlying assumptions of social proximity, we argue that the academic relations these students developed through graduate education can reduce the social distance between universities and firms, thus favoring collaborative research and technology transfer. Based on this argument, two hypotheses are presented to explain how the hiring of a former graduate student is associated with firms' collaboration decisions, constituting a driver of technology transfer. We empirically test these hypotheses with a new empirical strategy and use a novel and comprehensive dataset on university-industry linkages in Brazil. We find that approximately 40% of the collaborations were developed by firms with 'socially proximate' research groups, i.e., those hosted by universities where one or more firm employees attended graduate education. The estimates suggest that if a research group is socially proximate to a firm, the latter is more likely to choose this research group to partner with (relative odds approximately 2.5 times higher) and to engage in collaboration with (odds ratio more than 8 times higher). These results suggest new approaches for policy support to these partnerships, using academic relations as a lever to new collaborative projects.

Keywords University-firm collaborations · Graduate education · Social proximity · Conditional logit

JEL classification I23 · O30 · O31

✉ Daniel Gama e Colombo
dcolombo@alumni.usp.br

[1] Institute of Economics of the University of Campinas, Campinas, Brazil

1 Introduction

Graduate education has been growing steadily in recent decades (Nerad & Evans, 2014).[1] As the most advanced level of academic training, it provides students not only with deep knowledge and analytical skills but also with opportunities to interact with and develop relations with other researchers. Such relations constitute important social bonds, and they can help students foster new projects and partnerships after they graduate. Although the importance of such ties for collaborative research is acknowledged in the literature (Ponomariov, 2009), the measurement of their contribution is scarce, which indicates the need for novel empirical analyses on the topic.

Scientific partnerships are an important piece of the strategy for creating and transferring technology from universities to firms (Muscio, 2010), and they have become an important source of productivity enhancement for firms in different fields and sectors in recent decades (Bozeman et al., 2013; Colombo et al., 2021). Doctoral students and graduates can be a major asset for collaborations (Ponomariov, 2009), and for this reason, innovation policies have stressed the importance of doctoral education for collaborative arrangements (Thune, 2009). Firms hiring such professionals not only increase their ability to evaluate, assimilate and exploit external knowledge (Cohen & Levinthal, 1990) but also obtain access to the 'academic relations' of these individuals (Vinding, 2004), interpreted herein as the social ties developed with advisors, professors, peers and other members of their universities.

The existing research on this subject is limited and nonsystematic (Thune, 2009), and the related empirical evidence is scarce (Santos et al., 2020). In particular, there is still room to analyze these personal ties of former graduate students from the lens of knowledge transfer from academic research to industrial innovation (Granovetter, 2005; Mansfield, 1995). This literature has considered different factors as determinants of social proximity (Huber, 2012), including firms founders relations (Colombo et al., 2021) and 'employee-driven relations' of those who attended college at local universities (Drejer & Østergaard, 2017; Østergaard, 2009). However, social ties developed during doctorate training should also be important for scientific collaborations. This configures a gap in the existing literature and a promising research agenda that can provide new insights into the understanding of collaborative research.

This paper aims to fill this gap by investigating how the academic relations of former graduate students contribute to scientific collaboration between universities and firms by reducing the social distance between universities and firms. We explain how and test whether hiring a former graduate student is associated with the likelihood of a firm collaborating with a research group belonging to the employee's graduate university. The proposed hypotheses are tested empirically using firm-level data on university-industry linkages in Brazil. The results confirm that academic relations are significant predictors of collaborations, thus suggesting that they constitute an important component of social proximity between partners. The results also indicate that the magnitude of this association varies substantially according to the knowledge area.

Our contributions to the existing literature are threefold. The theoretical contribution is the analysis of academic relations in light of the underlying assumptions of social

[1] Herein interpreted as referring exclusively to master's and PhD programs, which constitute '*stricto* sensu' graduate education in Brazil, that is, the only programs that grant an academic degree, not including any other programs that award a nonacademic certificate (such as professional training).

proximity, using such a framework to explain how such relations constitute drivers of technology transfer and collaborative research. The empirical contribution is to test and measure the importance of such relations using a novel and comprehensive dataset and applying a new empirical strategy that models firms' decisions in two steps (the choice of partner and the decision to collaborate). The third contribution is to estimate the parameters of the model for each knowledge area separately. This addresses a major gap in the literature, as most empirical studies limit the analysis to a specific knowledge area or sector (Broekel & Hartog, 2013), although there are sound arguments to suspect that scientific disciplines can work as 'potential moderators' of proximity dimensions (Rybnicek & Königsgruber, 2019).

Our empirical analysis is based on Brazilian data. The landscape of innovation and university-firm collaboration in the country has been extensively discussed in previous studies (Mazzucato & Penna, 2016). Brazil has a low but heterogeneous innovation base, with a small group of excellent research centers. University-firm collaborations for innovation projects are highly concentrated in the southern part of the country (Garcia et al., 2015), with public support focused on low and medium-tech industries (Freitas et al., 2013).

The second section following this introduction briefly discusses the main developments and the state of the art of the relevant literature on university-firm partnerships and social proximity. The third section presents the main arguments and research hypotheses, which are tested using the data presented in the fourth section and according to the empirical strategy described in the fifth section. The sixth section presents and discusses the findings of the empirical analysis, and the final section summarizes the main results and suggests potential research topics for future studies.

2 Literature review

Universities have been increasingly engaging in collaboration with industry (OECD, 2019). However, academia and industry operate in distinct institutional environments that are characterized by norms and incentives that can conflict with one another (Partha & David, 1994). Research in universities and public research institutes can differ from that carried out by firms in terms of 'basicness', scope and impact (OECD, 2019). On the one hand, academic researchers are keen to generate new knowledge and publish their results in high-impact journals; on the other hand, industrial scientists and engineers are more interested in producing applied knowledge in the form of products and manufacturing processes.

These differences give rise to barriers that can hinder collaboration between these organizations. Due to orientation barriers, academic researchers can be reluctant to engage with industry because collaboration can pose potential dilemmas for scientific research (Perkmann et al., 2021). The motivation of academic researchers to collaborate with firms is mostly driven by the aim of promoting their own research agenda. In this way, collaborative projects with the private sector should be complementary to the norms of open science, and they require a better understanding of the context in which basic research is applied so that these researchers can access resources and skills that are not available in universities (Tartari et al., 2012). University-firm collaboration can also be hampered due to complementarity barriers, defined as the lack of complementarity between industry-related scientific activities and academic research.

The evolution of knowledge networks that comprise university-firm collaboration has received growing attention in the economic geography literature (Broekel, 2015; Ter Wal & Boschma, 2009). Recent studies have tried not only to explain the underlying dynamics of network evolution but also to understand to what extent geographical proximity is important to the establishment of collaboration ties among different partners.

Geographical proximity facilitates linkages due to the existence of mechanisms such as frequent interactions and face-to-face contacts. Firms often prefer to collaborate with proximal universities because they can reduce the costs of partnerships (D'Este et al., 2013), capture geographically bounded knowledge spillovers and understand local researchers' projects and activities due to their social ties (Drejer & Østergaard, 2017). However, other dimensions of proximity can foster interactive learning among partners. Cognitive proximity is related to the level of overlap in two actors' knowledge bases (Nooteboom et al., 2007), and there is evidence that it is important for increasing the frequency of collaborations (Muscio, 2013; Muscio & Pozzali, 2013). Institutional proximity indicates the degree to which two institutions are subject to the same institutional framework, background, and systems of rewards and values (Broekel, 2015; Ponds et al., 2007). Organizational proximity refers to the degree of strategic interdependence or control induced by the link between partners, such as the one shared by firms belonging to the same corporate group (Balland, 2012).

Within such a framework, social proximity measures the strength of interpersonal linkages or the extent to which individuals know each other and interact in personal or professional contexts (Huber, 2012). It describes agents' social embeddedness in terms of friendship, kinship and experiences (Granovetter, 1985). It also underlines the role of trust, which can be positively influenced by social proximity (Broekel, 2015; Nooteboom, 2002). The main argument is that strong, trust-based ties facilitate knowledge sharing and interactive learning (Gertler, 2003; Huber, 2012). Empirical studies have shown that social proximity increases the likelihood of linkages among actors (Cassi & Plunket, 2014; Huber, 2012).

The importance of graduate education for social proximity has not been properly addressed. However, former graduate students should also play an important role in the formation of knowledge networks (Ponomariov, 2009), not only because they add to the firm's absorptive capacity (Garcia-Quevedo et al., 2012) but also because of their social relations within the scientific community. Such relations are relevant, as they are mostly nurtured through proper scientific activities such as the development of research projects, publications, professional societies and attendance at congresses and professional meetings (Roach & Sauermann, 2010). As a result, these 'linked scientists' end up having personal relations in (and valued by) both the academic and private sectors, putting them in an ideal position to function as nodes of knowledge networks (Lam, 2005) or as 'bridge builders' between industry and academia (Thune, 2010), helping to overcome orientation and complementarity barriers.

There are different reasons why the social relations of these students and graduates may improve social proximity and help foster collaborative research and development (R&D). First, the 'relational capital' built during graduate training signals trust and respect (Attia, 2015; Canhoto et al., 2016), which discourages malfeasance and opportunistic behavior, smooths negotiations and constitutes a driver of partnerships between the private sector and the scientific community. Second, graduate degree holders are likely to share a common language with their university peers and professors, which can help them overcome information barriers (Baba et al., 2010; Gawel, 2014). Additionally, collaborating universities and firms face their costs of bridging their differences in norms and modes of operation (Colombo et al., 2021), and workers with graduate degrees can help to reduce this cultural gap, balancing different priorities, goals and timing among the organizations involved (Barnes et al., 2002). Finally, employees with graduate training can also strengthen firms' commitment to collaboration, increasing the willingness to allocate effort and resources to these projects (Attia, 2015).

The empirical literature on the importance of these social relations is still scarce, but the existing evidence suggests that doctoral students' research can positively influence knowledge transfer and collaboration (Santos et al., 2020). On the university side, scientists that interact with graduate students have been found to be more likely to collaborate with the industrial sector (Ponomariov, 2009), and there is also evidence that graduate students are important channels linking applied researchers and professors to the industry (Balconi & Laboranti, 2006).

The arguments and evidence discussed in this section suggest that incorporating the social relations of former graduate students within the economic geography framework can provide both relevant insights to understand university-firm collaborations and policy options to foster them. In the next section, we present a set of hypotheses aimed at filling this gap, using social proximity as the conceptual building block of the analysis.

3 Research hypotheses

We argue that professionals who have attended graduate education can act as catalysts of partnerships and technology transfer between institutions because they are an important source of knowledge and for their academic relations within the university where they pursued graduate-level studies.

We argue that a firm employing a former graduate student obtains access to his or her academic relations. Such employees serve as a link, increasing the social proximity between the employer organization and the university, thus affecting its incentives, costs and expected returns of collaboration (Colombo et al., 2021). Based on the theoretical arguments described in the previous section, we present two hypotheses regarding how employing a former graduate student can be associated with the likelihood of collaboration between universities and firms.

We first consider the choice of a research group or university with which to collaborate. We expect that the social proximity emerging from academic relations can reduce the perceived risks related to the partner's innovation capabilities, opportunistic behavior or lack of engagement. The former student in the workforce strengthens trust and commitment between partners (Attia, 2015; Canhoto et al., 2016; Teirlinck & Spithoven, 2013) because of previous interactions, long-term relations and the likelihood of future collaborations. Such employees are also in a better position to forecast the potential results of the collaboration and reduce orientation and complementarity barriers, as they are likely to have a better understanding of the skills and knowledge of the research group and of the institutional framework and culture of the university. Therefore, proximity can facilitate negotiations for knowledge transfer, the allocation of resources, and the division of project results. Consequently, academic relations reduce uncertainty and add value to the expected return of collaborative projects, which leads to the first hypothesis:

Hypothesis 1: Firms are more likely to choose a research group to collaborate with if one or more of its employees have attended graduate education at the group's host university.

A much less studied topic in this literature is the investigation of factors associated with the firm's decision to engage or not engage in collaboration with any university. However, this constitutes a promising research agenda to be explored, as the few existing studies on this topic have found evidence of the correlation of distinct factors with such decisions (Drejer & Østergaard, 2017; Maietta, 2015). The social proximity induced by academic relations is a relevant point to be considered in this agenda. When firms hire a former graduate student, they improve their social proximity with the scientific community of the employee's university by reducing the cultural and language gap, mitigating potential barriers and information asymmetries, and strengthening trust and commitment. Therefore, we expect academic relations to be associated with the decision to collaborate or not, as suggested by *Hypothesis 2*.

Hypothesis 2: Firms are more likely to engage in collaboration if they employ one or more former graduate students from the university that hosts the research group that is the most likely partner.

4 Data, sample and variables

To test the research hypotheses, we use a novel and rich dataset that comprises confidential microdata on a large number of firms in Brazil at the employee level, with information on their collaboration with academic research groups and the employment of former graduate students.[2] For the empirical analysis, we interpret as "firms" not only commercial firms but also other organizations operating under private law (private nonprofit organizations, publicly held corporations, and public organizations operating under private law).

We assemble the dataset by merging information from three databases: (i) the '2016 Census of Research Groups' (CNPQ, 2016); (ii) employment data from the 'Annual Social Information Report – RAIS' of the Ministry of Economics (2015); and (iii) the database of Brazilian graduate students (CAPES, 2017). The sample used in the study consisted of local units of firms with at least one employee with a higher education degree in 2015 and their collaborations with research groups within Brazilian universities in 2016. The employees analyzed to assess social proximity are those who ever enrolled in a Master's or PhD program from 1996 to 2015. The main fields of education of the research groups were used to classify them into different knowledge areas (UNESCO-UIS, 2015). Information on collaborations in the CNPQ (2016) database is self-reported by research groups, which provides a broader and more adequate indicator of such partnerships, as it is not publication-based and gives the respondents flexibility to decide whether an interaction with a firm can be considered a collaboration (Bozeman & Corley, 2004; Bozeman et al., 2013).

The main variable of interest is social proximity, represented by a dummy indicating whether the firm employed one or more former graduate students from the university that hosts the academic research group. Geographical proximity is measured by its inverse, i.e., the distance (in 100 km) between the cities of each part of the dyad, following previous papers that also used continuous variables to study or control for this factor (D'Este et al., 2013; Drejer & Østergaard, 2017; Garcia et al., 2018). The institutional proximity dimension is measured by a dummy indicating whether the host university is private, such that it is subject to a similar set of rules and legal statutes as the firm. And cognitive proximity is represented by the narrow knowledge area[3] within which firms search for a research

[2] Access to confidential data was granted by the Brazilian Institute of Educational Research (Inep) for purposes of this research.

[3] We use the 'main field of education' of each research group (CNPQ, 2016) as the narrow knowledge area.

partner. We also control for attributes of both the research group and the host university, and we add state dummies to indicate the location.

In the second stage, features of the firm are also included as explanatory variables, including sectoral dummies (ISIC, 2-digit level). We proxy the absorptive capacity of firms using both the share of employees with higher education (following Garcia et al., 2018 and Drejer & Østergaard, 2017) and the share with a graduate degree. In light of the structure and updating procedures of the databases used in the analysis, we assume that collaboration decisions in year '*t*' are based on attributes of the research groups and host universities in the same period and on features of the firms' and employees' educational attainment observed in period '*t -1*'.

All variables used in the study are listed in Table 1. For comparison purposes, descriptive statistics are presented for the entire sample and for only firms that collaborated.

Table 2 presents the distribution of collaborations and collaborating entities per knowledge area, and it provides a first idea of the importance academic relations and social proximity can have on the formation of these links, in favor of the research hypotheses. Approximately 40% of the collaborations were developed by firms with a research group from a host university where one or more firm employees attended graduate education. Such a proportion is higher in 'Education', 'Social Sciences, Journalism and Education' and 'Health and Welfare'. Table 2 also shows that firms developed, on average, more than two collaborations in the relevant period.

5 Empirical strategy

The objective of the empirical analysis is to investigate whether hiring a former graduate student is associated with the firm's decision to collaborate with a research group hosted by the employee's university. We use a multiple-stage model to describe the firm's decision-making process for collaborating with a research group, following De Fuentes and Dutrénit (2012) and Laursen et al. (2011). The model used herein was originally developed by Long (2004) and Skinner (2019) to explain higher education choices. We adapted this framework to the context of university-firm collaboration, an original approach not considered in previous studies.

Each collaboration available to a firm is considered 'a package containing various attributes' (Long, 2004), including the dimensions of proximity. Firms compare the expected returns of all potential choices of collaboration, including the option of not partnering with any research group. Although one can observe only the final decision made by the firm (i.e., to collaborate with a specific group or to not collaborate at all), this choice can be divided into two stages, and each one tests one of the above research hypotheses. First, the firm considers all potential research groups available for collaboration, identifying the one with the highest expected net result. In the second stage, it compares such choice with the option of not collaborating, deciding the best course of action.

5.1 The first stage: choice of partner

In the first stage, firms search among all potential partners to identify the most rewarding collaboration, considering the expected costs, returns and risks of each choice. To make this decision, they consider features of the research group and its host university, along

Table 1 Descriptive statistics of the variables used in the empirical analysis. *Source*: prepared by the authors based on CAPES (2017), CNPQ (2016), and Ministry of Economics (2015)

Variables	(1) All firms Mean (Std. dev.)	(2) Only collaborative firms Mean (Std. dev.)
Collaboration with a research group (dummy)	0.001 (0.04)	1 (0)
Number of collaborations		2.5 (10.42)
Features of the firm		
Size (number of employees)	12.62 (87.27)	359.85 (954.48)
Organization incorporated as a commercial firm (dummy)	0.92 (0.27)	0.72 (0.45)
Absorptive capacity: percentage of employees with an undergraduate degree	0.09 (0.22)	0.41 (0.32)
Absorptive capacity: percentage of employees with a graduate degree[a]	0.001 (0.02)	0.04 (0.1)
Attributes of the research group and host university[b]		
Host university incorporated as a commercial enterprise (dummy)		0.12 (0.33)
Age of the research group		14.50 (9.57)
Number of researchers in the research group[c]		12.54 (8.96)
Number of private partners of the research group		7.15 (11.86)
Proximity factors[b]		
Geographical distance (per 100 km)[d]		3.05 (4.96)
Institutional (private host university)		0.203 (0.36)
Social (dummy for employment of a former graduate student from the host university)		0.28 (0.40)
Number of former graduate students from the host university employed by the firm		4.23 (19.21)
No. of obs.[e]	2,247,423	3,225

[a]Master's or PhD degree

[b]Each collaborating firm considered as a single observation. In case of multiple collaborations, the mean value per firm is considered

[c]Not considered students, external members and technical staff

[d]Distance between municipalities

[e]Local units of firms

Table 2 Number of firms, research groups, collaborations and social proximity per knowledge area. *Source*: CAPES (2017), CNPq (2016), and Ministry of Economics (2015)

Broad field of education and training	(1) Firms (freq.)	(2) Research groups (freq.)	(3) Average number of research groups per narrow knowledge area	(4) Collaborations (freq.)	(5) Average number of collaborations per firm	(6) Collaborations between social proximate organizations[a] (%)
Education	223	949	638.2	407	1.8	50.12
Arts and Humanities	159	738	106.3	246	1.5	43.50
Social Sciences, Journalism and Information	292	943	131.2	482	1.7	48.55
Business, Administration and Law	248	494	260.4	409	1.6	40.83
Natural Sciences, Mathematics and Statistics	771	2,730	246.0	1,551	2.0	45.39
Information and Communication Technologies	235	334	334.0	342	1.5	37.72
Engineering, Manufacturing and Construction	1,329	1,807	183.1	2,428	1.8	38.80
Agriculture, Forestry, Fisheries and Veterinary	705	1,224	336.4	1,360	1.9	32.21
Health and Welfare	482	1,647	312.1	823	1.7	47.02
Services	13	31	20.4	14	1.1	42.86
All areas[b]	3,225	10,897	261.8	8,062	2.5	41.16

[a]Firm employs a former graduate student from the host university (*social proximity dummy = 1*)

[b]Firms can collaborate in more than one knowledge area

with proximity factors, including the social dimension, represented by the academic relations of former graduate students (*Hypothesis 1*). To control for cognitive proximity, we follow Ponds et al. (2007) and assume that the search for a partner is limited to a narrow knowledge area such that all potential research groups available for collaboration present a (similar) small cognitive distance.

We model this stage as a probabilistic equation (Long, 2004; Skinner, 2019) to measure and test the association of different factors with this decision. The conditional probability (P_{ij}) that a random firm i will choose a research group j *(choice$_i$=j, where j=1, 2, … J)* as the one with the highest net expected result is (Greene, 2011)

$$P_{\mathrm{ij}} = \mathrm{Prob}\left(\mathrm{choice}_i = j | X_{\mathrm{ij}}\right) = \frac{\exp\left(\beta \prime x_{\mathrm{ij}}\right)}{\sum_{k=1}^{J} \exp\left(\beta \prime x_{\mathrm{ik}}\right)} \tag{1}$$

where X_{ij} is a vector of explanatory variables associated with the probability that a firm will choose a particular research group (the '*attributes of the research group and host university*' and '*proximity factors*' listed in Table 1), and β is the vector of parameters (to be estimated) that indicate the magnitude of the association. The estimation is based on the conditional logit or McFadden's discrete choice model (Greene, 2011), which is suitable for cases in which the decision-maker is faced with a great number of choices, as it exploits the variation of attributes and interaction terms (Long, 2004). This model also has the advantage of controlling for individual attributes of firms that do not need to be included as independent variables, as they are differenced out of the equation.

For this estimation, we use data from collaborative firms and research institutes, assuming that all firms have decided on the collaborations with the highest expected return. The initial dataset is expanded to cover all potential choices of collaborative firms, i.e.: for each actual collaboration, we pair the firm with all existing research groups within the narrow knowledge area of the actual partner in the original dataset. Each choice is considered separately, so if a firm collaborated more than once, the dataset is expanded for each collaboration to comprise all possible dyads. The resulting expanded database comprises 2,110,638 observations,[4] and it includes a dummy that informs the realized ties (the original choice of collaboration), which constitute the response variable of the probabilistic model.

5.2 The second stage: decision to collaborate

In the second stage, firms consider the expected costs and benefits of their best choice for collaboration (selected at the first stage), comparing it with the alternative of not engaging in collaboration. For analytical purposes, this stage encompasses not only the firm's decision but also other transaction costs, such as project negotiation, intellectual property issues, and any other obstacles that need to be overcome for collaboration to occur.

To investigate how proximity factors and academic relations are associated with the decision at this stage, we must include firms that did not collaborate in the dataset.

[4] The number of collaborations (column 4 of Table 2) multiplied by the number of available research groups in the respective narrow knowledge area (column 3 of Table 2).

However, in these cases, we do not have a chosen research group with which to calculate the proximity variables. To overcome this problem, we pair each firm with an estimated 'most likely partner' (Long, 2004; Skinner, 2019), as follows: first, a narrow knowledge area of interest is selected for each firm[5]; then, we expand the dataset to cover all potential collaboration choices for each firm (i.e., all research groups within the respective narrow knowledge area); third, we apply the β parameters estimated at the first stage (for each broad field) to all dyads and calculate the probability that each link is formed (P_{ij}), with no minimum probability cutoff; finally, the research group with highest P_{ij} is selected as the 'most likely partner' and paired with the respective firm, excluding all other potential choices. The most likely partner of firms that actually collaborated may not necessarily be the one in the original dataset.[6]

The response variable in the second stage is a dummy that represents the decision of firm i to engage in collaboration *(collaboration*$_i$*=1*) or not.

(collaboration$_i$*=0*). The decision is modeled as presented in Eq. 2, where P_{ij} is the probability that firm i decides to engage in collaboration, conditional on the 'most likely' partner chosen at the first stage being j. Z_{ij} is the set of explanatory variables associated with this decision listed in Table 1, which includes '*features of the firm', 'attributes of the research group and host university',* and '*proximity factors*'. The associations of different factors with the likelihood of the outcome are represented by the vector β, which is estimated using a standard logistic regression (Greene, 2011).

$$P_{\mathrm{ij}}\left(\mathrm{collaboration}_i = 1|\mathrm{choice}_i = j, Z_{\mathrm{ij}}\right) = \frac{\exp\left(\beta\prime z_{\mathrm{ij}}\right)}{1 + \exp\left(\beta\prime z_{\mathrm{ij}}\right)} \tag{2}$$

6 Results and discussion

Tables 3 and 4 present the estimated parameters for the first and second stages of the model. They are presented in odds ratios, which inform how a unit increase in the value of an explanatory variable (holding all others constant) is associated with a change in the relative odds of the outcome represented by the dependent variable (i.e., the odds of a firm engaging in collaboration divided by the odds of the firm not collaborating). The association is positive if the estimated odds ratio is greater than one, meaning that a higher value of the explanatory variable is correlated with a higher likelihood of collaboration. On the other hand, a negative association is represented by an odds ratio coefficient between zero and one.[7] The statistical significance of the parameters was assessed at a 0.05 significance threshold (a 95% confidence interval).

[5] For firms that collaborated, the narrow knowledge area of interest is the one of the actual partner research groups, while for firms that did not collaborate, we use the most frequent narrow knowledge area of partners of firms of the same sector (using the International Standard Industrial Classification – ISIC, 2-digit level).

[6] The similarities between the actual and estimated most-likely partners are discussed with the results of the empirical analysis.

[7] Therefore, e.g., the odds that a research group is chosen by a collaborative firm at the first stage is 1.7% higher than the odds of a younger group for each year of difference in age between them (column 1 of Table 3). On the other hand, commercial firms are 54.8% (*1 – 0.452*) less likely (i.e., lower relative odds) to engage in collaboration at the second stage than noncommercial firms (column 1 of Table 4).

Table 3 Estimated parameters of the first stage—choice of partner. Dependent variable: $choice_i$ (dummy for the research group that collaborated with each firm). *Source*: prepared by the authors based on CAPES (2017), CNPQ (2016), and Ministry of Economics (2015)

Independent variables	Knowledge Area[a]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All fields	Education	Arts and Humanities	Social Sciences, Journalism and Information	Business, Administration and Law	Natural Sciences, Mathematics and Statistics	Information and Communication Technologies	Engineering, Manufacturing and Construction	Agriculture, Forestry, Fisheries, and Veterinary	Health and Welfare
Proximity factors										
Geographical distance (100 km)	0.789*** (0.004)	0.779*** (0.020)	0.742*** (0.034)	0.791*** (0.017)	0.775*** (0.021)	0.779*** (0.010)	0.832*** (0.020)	0.795*** (0.009)	0.769*** (0.009)	0.774*** (0.016)
Institutional (private host university)	1.147*** (0.041)	1.348** (0.181)	1.398 (0.296)	1.399** (0.188)	1.106 (0.132)	1.289*** (0.118)	1.239 (0.193)	1.072 (0.071)	1.039 (0.148)	0.970 (0.107)
Social (dummy for employment of a former graduate student from the host university)	2.468*** (0.088)	2.373*** (0.412)	1.698** (0.365)	2.287*** (0.345)	1.587*** (0.245)	2.360*** (0.186)	3.176*** (0.537)	2.939*** (0.200)	1.890*** (0.168)	2.559*** (0.295)
Attributes of the research group and host university										
Host university incorporated as a commercial enterprise (dummy)	0.772*** (0.060)	1.499 (0.418)	0.473 (0.472)	1.535 (0.440)	1.194 (0.246)	1.288 (0.240)	1.378 (0.475)	0.677* (0.139)	1.069 (0.178)	1.359 (0.277)
Age of the research group	1.017*** (0.001)	1.001 (0.008)	1.004 (0.009)	1.014** (0.007)	1.021*** (0.008)	1.012*** (0.002)	1.008 (0.009)	1.021*** (0.002)	1.020*** (0.002)	1.017*** (0.004)

Table 3 (continued)

Independent variables	Knowledge Area[a]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All fields	Education	Arts and Humanities	Social Sciences, Journalism and Information	Business, Administration and Law	Natural Sciences, Mathematics and Statistics	Information and Communication Technologies	Engineering, Manufacturing and Construction	Agriculture, Forestry, Fisheries, and Veterinary	Health and Welfare
Number of researchers in the research group	1.011*** (0.001)	0.988** (0.006)	0.989 (0.007)	1.003 (0.005)	0.997 (0.006)	1.010*** (0.003)	1.003 (0.004)	1.015*** (0.002)	0.988*** (0.004)	1.011*** (0.003)
Number of private partners of the research group	1.070*** (0.002)	1.167*** (0.012)	1.134*** (0.018)	1.110*** (0.012)	1.143*** (0.013)	1.092*** (0.005)	1.137*** (0.013)	1.058*** (0.003)	1.137*** (0.006)	1.135*** (0.010)
State dummies[b]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chi-squared	7,630	19,344	17,922	1,876	25,464	2,881	36,935	5,264	2,075	4,748
Prob > chi2	0	0	0	0	0	0	0	0	0	0
Pseudo-R^2	0.180	0.212	0.206	0.207	0.174	0.183	0.168	0.201	0.177	0.190
Log-pseudo-likelihood	−35,213	−2,030	−867.8	−1,825	−1,875	−6,684	−1,653	−9,847	−6,264	−3,683
No. of Obs	2,110,638	259,765	26,146	63,239	106,484	381,580	114,228	444,637	457,442	256,831

[a]According to the classification of the ‘broad fields of education and training’ presented by UNESCO-UIS (2015). The ‘Services’ field was not estimated due to the small sample size

[b]Location of research groups

Conditional Logit Model. Odds ratios are reported. Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4 Estimated parameters of the second stage—decision to collaborate. Dependent variable: *collaboration*$_i$ (dummy indicating whether the firm engaged in collaboration). *Source*: prepared by the authors based on CAPES (2017), CNPQ (2016), and Ministry of Economics (2015)

Independent variables	Knowledge Area[a]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All fields	Education	Arts and Humanities[b]	Social Sciences, Journalism and Information	Business, Administration and Law	Natural Sciences, Mathematics and Statistics	Information and Communication Technologies	Engineering, Manufacturing and Construction	Agriculture, Forestry, Fisheries, and Veterinary	Health and Welfare
Proximity factors										
Geographical distance (100 km)	1.015	1.009	0.487	1.024	0.144**	1.082	0.924	0.986	1.205***	0.865**
	(0.009)	(0.049)	(0.225)	(0.076)	(0.110)	(0.091)	(0.159)	(0.025)	(0.052)	(0.050)
Institutional (private host university)	0.961	1.019		0.145***	7.038	1.627	9.232	0.261***	0.241***	0.505**
	(0.077)	(0.496)		(0.088)	(8.476)	(1.697)	(17.051)	(0.050)	(0.101)	(0.165)
Social (dummy for employment of a former graduate student from the host university)	8.607***	4.511***		6.785***	14.157***	20.944***	30.537***	8.049***	10.245***	17.360***
	(0.584)	(1.507)		(4.576)	(11.587)	(22.961)	(24.093)	(1.233)	(2.173)	(5.730)
Attributes of the research group and host university										
Host university incorporated as a commercial enterprise (dummy)	1.150	0.790		7.419**	0.005**	32.258		3.716***	3.246***	1.790
	(0.127)	(0.994)		(6.214)	(0.011)	(108.617)		(1.137)	(1.382)	(0.810)
Age of the research group	1.007***	1.032	2.491***	1.063*	1.145	0.986	1.019	0.969***	0.968***	0.970*
	(0.002)	(0.029)	(0.746)	(0.035)	(0.141)	(0.020)	(0.060)	(0.006)	(0.009)	(0.015)

Table 4 (continued)

Independent variables	Knowledge Area[a]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All fields	Education	Arts and Humanities[b]	Social Sciences, Journalism and Information	Business, Administration and Law	Natural Sciences, Mathematics and Statistics	Information and Communication Technologies	Engineering, Manufacturing and Construction	Agriculture, Forestry, Fisheries, and Veterinary	Health and Welfare
Number of researchers in the research group	0.984***	1.027		1.038**	1.197***	0.996	1.002	0.996	0.987	0.992
	(0.002)	(0.019)		(0.017)	(0.080)	(0.024)	(0.089)	(0.004)	(0.014)	(0.014)
Number of private partner of the research group	0.981***	1.027		1.176***	1.264**	0.989	1.254	0.996	0.862***	0.973
	(0.004)	(0.042)		(0.051)	(0.147)	(0.038)	(0.295)	(0.008)	(0.027)	(0.031)
Features of the firm										
Size (number of employees)	1.001***	1.001*	1.117**	1.001***	1.004***	1.000**	1.001***	1.001***	1.001***	1.001***
	(0.000)	(0.000)	(0.049)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Organization incorporated as a commercial firm (dummy)	0.452***	0.042***		0.002***	0.000***	0.017***	0.126*	1.173	4.162***	0.495***
	(0.037)	(0.024)		(0.001)	(0.000)	(0.008)	(0.149)	(0.676)	(0.780)	(0.114)

Table 4 (continued)

Independent variables	Knowledge Area[a]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All fields	Education	Arts and Humanities[b]	Social Sciences, Journalism and Information	Business, Administration and Law	Natural Sciences, Mathematics and Statistics	Information and Communication Technologies	Engineering, Manufacturing and Construction	Agriculture, Forestry, Fisheries, and Veterinary	Health and Welfare
Absorptive capacity: share of employees with an undergraduate degree	7.048***	22.179***		7.973***	55.444***	4.742**	8.099***	8.143***	9.061***	3.665***
	(0.380)	(9.622)		(3.319)	(78.539)	(3.133)	(3.213)	(1.036)	(1.276)	(0.947)
Absorptive capacity: share of employees with a graduate degree	6.236***	17.178***		0.302	27.312*	0.047	0.805	2.676*	0.279	2.438
	(0.855)	(7.331)		(0.429)	(52.798)	(0.176)	(0.811)	(1.589)	(0.252)	(2.961)
Constant	0.000***	0.033***	39.567	0.234	25.982	5.709	0.000***	0.000***	0.022***	0.000***
	(0.000)	(0.041)	(233.819)	(0.256)	(83.778)	(12.841)	(0.000)	(0.000)	(0.016)	(0.000)
State dummies[b]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sector dummies (ISIC 2-digit)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4 (continued)

Independent variables	Knowledge Area[a]									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All fields	Education	Arts and Humanities[b]	Social Sciences, Journalism and Information	Business, Administration and Law	Natural Sciences, Mathematics and Statistics	Information and Communication Technologies	Engineering, Manufacturing and Construction	Agriculture, Forestry, Fisheries, and Veterinary	Health and Welfare
Percentage of firms that collaborated (*collaboration*$_i$=*1*)	0.14%	0.09%	1.84%	0.23%	0.28%	0.19%	0.14%	0.08%	0.04%	0.11%
Chi-squared	14,874	556.6	236.5	440.3	185.9	1,129	286.0	3157	2,674	1,176
Prob > chi2	0	0	0	0	0	0	0	0	0	0
Pseudo-R^2	0.321	0.394	0.343	0.716	0.819	0.787	0.248	0.305	0.304	0.311
Log-pseudo-likelihood	−16,458	−538.4	−16.38	−148.5	−52.14	−122.3	−301.4	−3,605	−2,093	−859.1
No. of obs	2,247,423	122,728	272	32,097	14,985	42,478	37,397	827,943	948,538	148,608

[a]According to the classification presented in UNESCO-UIS (2015)

[b]Location of firms. Logit Model. Odds ratios are reported. Robust standard errors in parentheses

***$p<0.01$, **$p<0.05$, *$p<0.1$

The estimates for the first stage (Table 3) provide strong empirical support for *Hypothesis 1*. For all knowledge areas (and both for the entire sample), the estimated social proximity coefficients are positive and statistically significant, which suggests that firms are more likely to collaborate with a research group if it is hosted by a university at which one or more of their employees pursued graduate-level studies. Considering the entire sample (column 1), the relative odds are 2.5 times that of groups with no social proximity. Although the magnitude of parameters varies per knowledge area, the likelihood is at least 69.8% higher (odds ratio found for 'Business, Administration and Law').

The coefficients estimated for the first stage were used to predict the most likely partner for each firm in the dataset. Descriptive statistics and the similarities between actual and estimated most likely partners for collaborative firms are presented in Tables A1 and A2 of the Online Appendix. Most attributes and proximity factors for both partners present similar values,[8] suggesting that predicted partners have similar features to those actually chosen by collaborative firms in terms of observable features, thus providing grounds for accepting the prediction model.

The results for the second stage in Table 4 show that social proximity is positively associated with the decision to engage in collaboration, as suggested by *Hypothesis 2*. Considering the whole sample, firms are more than eight times as likely to collaborate if one or more of their employees have attended graduate education at the university that hosts the most likely partner. We also find a positive and significant association for most knowledge areas.

The empirical results support the hypotheses regarding the association between employing a former graduate student and the likelihood of collaboration presented in Sect. 3. The importance of academic relations can be explained by the social dimension of proximity, based on the ideas of trust, common language, culture, and commitment (Rybnicek & Königsgruber, 2019), and how these factors affect the expected costs and returns arising from collaborations. These are novel findings on the role of social proximity in fostering university-firm collaboration, since we incorporate the value of academic relations to scientific collaborations. They confirm that former graduate students play an important role as 'linked scientists' or nodes of knowledge networks (Lam, 2005) and that their interpersonal linkages are representative of the social dimension of proximity (Huber, 2012). As graduate students leave academia to work in private organizations, their personal and professional connections within the scientific community bring organizations closer.

Confirming the theoretical expectations, academic relations of former graduate students facilitate communication and negotiation between firms and academic research groups, adding value to the expected return of the collaborative initiative (Colombo et al., 2021). As social proximity improves the prospects of successful and profitable collaboration, firms are prone to collaborating with socially proximate research groups (*Hypothesis 1*), in addition to becoming more inclined to pursue this strategy to develop new technologies and find new market opportunities (*Hypothesis 2*).

Our findings can also be construed as evidence of the role of former graduate students in building bridges that enable collaboration between academic research and firms, mitigating potential orientation and complementarity barriers (Partha & David, 1994; Tartari et al., 2012). These professionals are in general socially closer to academic researchers,

[8] The social proximity variable presents the same value in 80% of the cases; the values for institutional proximity matches 71% of the times; and the average difference of geographical distance is only 0.72 (14% of the standard deviation presented in Table 1).

being in a better position to build a strong relationship between academic researchers and the R&D staff of the collaborative firm. In this way, former graduate students can help to manage differences in alignment and incentives between academics and industrial scientists and engineers.

Another important result is that the share of firms graduate personnel is also an important predictor of collaboration, as firms are more than 6 times as likely to collaborate per additional share of employees with a graduate degree in their workforce (considering the entire sample). These estimates may be construed as evidence of another channel through which graduate degree personnel may contribute to research collaboration (in addition to their academic relations) and indirectly to innovation. The results are in line with the argument that personnel with graduate degrees enhance the absorptive capacity of firms (Balland, 2012; Garcia-Quevedo et al., 2012; Giuliani & Bell, 2005), constituting ‘vessels of knowledge transfer’ (Thune, 2009). These highly qualified employees enhance firms’ ability to assimilate knowledge and generate new technology and henceforth to extract value from research collaborations, making them more prone (and therefore more likely) to collaborate.

Our findings also confirm that scientific collaboration is spatially concentrated, as concluded by nearly all previous studies on the subject (Hinzmann et al., 2019; Rybnicek & Königsgruber, 2019). In the first stage, firms are approximately 21% less likely to choose a research group to partner with for each 100 km of distance between their cities (whole sample estimate in column 1). Estimated parameters for different knowledge areas are all statistically significant and similar, with odds ratios ranging from 0.74 to 0.83. This finding can be explained by the typical forms of interaction between actors at the local level, such as face-to-face contact and frequent communications. Through face-to-face contact, agents can more easily solve conflicts and align expectations, with positive effects for the interactive learning process. The search for common interests among actors becomes an easier task, even when different organizations with different incentive systems are involved. Additionally, local interactions can also reduce the costs of knowledge transfer across larger distances. (Fitjar & Gjelsvik, 2018; Ponds et al., 2007).

While our empirical analysis is restricted to Brazil, we believe that the arguments and findings presented in this paper are general enough to be applied to other contexts for three main reasons. First, firms both in Brazil and in other economies are seeking new sources of technological and scientific knowledge to support innovation. In the context of new knowledge-intensive technologies (often associated with so-called ‘Industry 4.0’), firms are continuously pushed to intensify their search for new technological knowledge, while universities continue to be a major source of it. Second, building channels of interaction between universities and firms is a growing challenge not only for firms and universities but also for policy-makers in any country, and our results suggest important insights that contribute to this goal. Finally, our theoretical arguments do not rely on any of the previously mentioned characteristics of the Brazilian industrial and innovation landscape, and our empirical results do not seem to be critically dependent on such features.

In recent decades, university-firm collaboration has entered the policy agenda as a strategy for fostering technology transfer (Muscio, 2010). The findings presented herein support policies aimed at promoting university-firm collaboration by boosting and exploiting the social networks of former graduate students. Suggested measures currently in place in different countries include tax breaks for firms hiring master’s and PhD graduates, financial support and the promotion of ‘revolving door’ mobility of industrial researchers (temporary placement in universities), joint supervision and co-financing, and ‘soft instruments’ to raise researchers’ awareness of entrepreneurship (OECD, 2019; Thune, 2009). These instruments, however, are mostly recent and restricted to a few initiatives in industrialized countries, thus requiring further investigations on their actual impact and applicability to less developed industrial economies. In addition, such measures should be considered within a broader policy strategy that also addresses other challenges of university-firm collaboration, such as financial and regulatory constraints to knowledge transfer and lack of awareness of the importance of collaborations and entrepreneurship by university staff.

This empirical analysis has limitations that must be considered when interpreting the results. First, it relies on data available in the original databases, following a 'complete case analysis' approach to deal with any missing data (Hughes et al., 2019). Although we expect our dataset to be representative of the collaborations between universities and firms in the country, we cannot ensure that nonreported or nonavailable data are 'missing at random'. In addition, the conditional logit used to estimate the first-stage parameters is based on the strong assumption of independence of irrelevant alternatives (IIA), although there are good arguments to maintain that this should not pose a threat of bias in this case (Skinner, 2019), i.e., completeness of the choice set; independence of the negotiations and transactions for different research groups; and the differences and unique features of each group and university.

The results of the second stage present additional limitations that are worth noting. First, the low proportion of correct predictions of the most likely partner (as presented in Table A1) suggests that there is room for improvement of the prediction model. Second, the small proportion of firms that actually collaborated (as reported in Table 4) indicates the possibility of a 'rare events' problem in the results (Leitgöb, 2020; Van der Paal, 2014), although the pseudo-R^2 of the estimates are all above 0.2, indicating a good fit of the model (Hemmert et al., 2018; McFadden, 1979).[9] Finally, the social proximity coefficient presents very high values (odds ratio above 20) for two knowledge areas,[10] suggesting the importance of investigating these collaborations more carefully. Mitigating these problems and confirming the findings of the second stage constitute important topics to be addressed by future research.

Finally, in addition to these limitations, this study does not aim to demonstrate a causal effect between the explanatory and dependent variables. Accordingly, the empirical results only confirm that academic relations are associated with collaboration decisions, thus constituting significant predictors of such partnerships. Proving the effect or the channels through which social proximity actually affects collaboration is a promising research agenda that falls outside the scope of this paper.

7 Concluding remarks

Universities and governments are striving to foster scientific collaborations to enhance technological development and transfer, providing benefits for both partners and for the economy. Alumni of graduate programs are in a privileged position to foster these partnerships, as they are a point of contact between the scientific and industrial communities, with

[9] 'Rare events' estimators may create additional problems by overcorrecting (Leitgöb, 2013) or introducing other biases in the estimates (Elgmati et al., 2015; Puhr et al., 2017). For such reasons, we followed the previous studies (Long, 2004; Skinner, 2019) that have used the standard logistic regression to estimate the second stage.

[10] 'Natural Sciences, Mathematics and Statistics' and 'Information and Communication Technologies'.

personal relations on both sides that can help to reduce uncertainties, build trust, and set goals and objectives and thereby reach agreements that are beneficial from both the scientific and commercial perspectives.

This paper discusses how the academic relations of former graduate students improve social proximity between firms and universities. Using a two-step model estimated with a novel database, we find that if a research group is hosted by a university in which one or more employees of a firm attended graduate-level education, the organization is more likely to choose this group to partner with (relative odds approximately 2.5 times higher) and to engage in collaboration with (odds ratio more than 8 times higher). Positive and statistically significant associations are found in both stages for the entire sample and for nearly all knowledge areas. The magnitude of the association varies depending on the knowledge area, in support of the argument that scientific disciplines work as 'moderators' of proximity factors. These results constitute the main contributions of the paper to the understanding of university-firm collaboration and to the creation and transfer of technology between them.

The results of this analysis contribute to the literature not only by adding to the (small) existing evidence on the importance of former graduate students to university-firm collaboration but also by building and interpreting this evidence based on the framework of economic geography and considering the different dimensions of knowledge networks (Broekel, 2015; Ter Wal & Boschma, 2009). From this perspective, the importance of academic relations is explained through the social dimension of proximity (Rybnicek & Königsgruber, 2019). The empirical strategy constitutes an additional contribution, as it allows us to test and confirm that social proximity and academic relations are associated both with the choice of the academic partner and with the decision to engage in collaboration.

The analysis also points to future research questions to improve our understanding of the connections between graduate education and scientific collaboration. The parameters estimated for individual knowledge areas suggest that academic relations may play different roles in each knowledge area, pointing to the need for specific studies dedicated to each one. Additionally, as previous studies have used other measures of social proximity (Broekel, 2015; Cassi & Plunket, 2014; Drejer & Østergaard, 2017; Hong & Su, 2013; Østergaard, 2009; Petruzzelli, 2011), an empirical analysis considering all these variables may provide a clearer picture of how these different networks are related and how they predict university-firm collaboration. It would also be important to replicate this empirical investigation using data from other countries to confirm that the arguments and findings presented herein can be generalized to other economies and industrial contexts. Finally, modeling academic relations using the number or share of former students in the firm's workforce (instead of the dummy variable used herein) can be a possible extension of this work, providing insights into how the interaction between multiple former students can contribute to the formation of ties.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10961-021-09881-2.

Declarations

Conflict of interest We have no conflict of interest to declare.

Data availability The data that support the findings of this study are confidential and their secrecy is protected by the law. The data were accessed for purposes of this study, under license granted by the Brazilian National Institute of Educational Studies and Research - INEP (Process 23036.006119/2019-01).

Code availability Code for data cleaning and analysis is available upon request to the authors.

References

Attia, A. M. (2015). National innovation systems in developing countries: Barriers to university–industry collaboration in Egypt. *International Journal of Technology Management & Sustainable Development, 14*(2), 113–124. https://doi.org/10.1386/tmsd.14.2.113_1

Baba, Y., Yarime, M., & Shichijo, N. (2010). Sources of success in advanced materials innovation: The role of" core researchers" in university–industry collaboration in Japan. *International Journal of Innovation Management, 14*(02), 201–219. https://doi.org/10.1142/S1363919610002611

Balconi, M., & Laboranti, A. (2006). University–industry interactions in applied research: The case of microelectronics. *Research Policy, 35*(10), 1616–1630. https://doi.org/10.1016/j.respol.2006.09.018

Balland, P.-A. (2012). Proximity and the evolution of collaboration networks: Evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies, 46*(6), 741–756. https://doi.org/10.1080/00343404.2010.529121

Barnes, T., Pashby, I., & Gibbons, A. (2002). Effective university–industry interaction: A multi-case evaluation of collaborative r&d projects. *European Management Journal, 20*(3), 272–285. https://doi.org/10.1016/S0263-2373(02)00044-0

Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy, 33*(4), 599–616. https://doi.org/10.1016/j.respol.2004.01.008

Bozeman, B., Fay, D., & Slade, C. P. (2013). Research collaboration in universities and academic entrepreneurship: The-state-of-the-art. *The Journal of Technology Transfer, 38*(1), 1–67. https://doi.org/10.1007/s10961-012-9281-8

Broekel, T. (2015). The co-evolution of proximities–a network level study. *Regional Studies, 49*(6), 921–935. https://doi.org/10.1080/00343404.2014.1001732

Broekel, T., & Hartog, M. (2013). Explaining the structure of inter-organizational networks using exponential random graph models. *Industry and Innovation, 20*(3), 277–295. https://doi.org/10.1080/13662716.2013.791126

Canhoto, A. I., Quinton, S., Jackson, P., & Dibb, S. (2016). The co-production of value in digital, university–industry R&D collaborative projects. *Industrial Marketing Management, 56*, 86–96. https://doi.org/10.1016/j.indmarman.2016.03.010

CAPES. (2017). *Discentes da Pós-Graduação Stricto Sensu do Brasil (base de dados confidencial) [Graduate Students in Brazil - confidential dataset]* CAPES - Coordination for the Improvement of Higher Education Personnel.

Cassi, L., & Plunket, A. (2014). Proximity, network formation and inventive performance: In search of the proximity paradox. *The Annals of Regional Science, 53*(2), 395–422. https://doi.org/10.1007/s00168-014-0612-6

CNPQ. (2016). *Censo do Diretório dos Grupos de Pesquisa 2016 [Research Group Census Database]* CNPQ - Brazilian Council of of Scientific and Technological Development, Brasilia. http://lattes.cnpq.br/web/dgp/censo-atual.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128–152. https://doi.org/10.2307/2393553

Colombo, M. G., Guerini, M., Rossi-Lamastra, C., & Bonaccorsi, A. (2021). The "first match" between high-tech entrepreneurial ventures and universities: The role of founders' social ties. *The Journal of Technology Transfer*. https://doi.org/10.1007/s10961-021-09848-3

De Fuentes, C., & Dutrénit, G. (2012). Best channels of academia–industry interaction for long-term benefit. *Research Policy, 41*(9), 1666–1682. https://doi.org/10.1016/j.respol.2012.03.026

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. https://doi.org/10.1093/jeg/lbs010

Drejer, I., & Østergaard, C. R. (2017). Exploring determinants of firms' collaboration with specific universities: Employee-driven relations and geographical proximity. *Regional Studies, 51*(8), 1192–1205. https://doi.org/10.1080/00343404.2017.1281389

Elgmati, E., Fiaccone, R. L., Henderson, R., & Matthews, J. N. S. (2015). Penalised logistic regression and dynamic prediction for discrete-time recurrent event data. *Lifetime Data Analysis, 21*(4), 542–560. https://doi.org/10.1007/s10985-015-9321-4

Fitjar, R. D., & Gjelsvik, M. (2018). Why do firms collaborate with local universities? *Regional Studies, 52*(11), 1525–1536. https://doi.org/10.1080/00343404.2017.1413237

Freitas, I. M. B., Marques, R. A., & e Silva, E. M. D. P. (2013). University–industry collaboration and innovation in emergent and mature industries in new industrialized countries. *Research Policy, 42*(2), 443–453. https://doi.org/10.1016/j.respol.2012.06.006

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. https://doi.org/10.1080/21681376.2015.1099464

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. https://doi.org/10.1080/23792949.2018.1484669

Garcia-Quevedo, J., Mas-Verdú, F., & Polo-Otero, J. (2012). Which firms want PhDs? An analysis of the determinants of the demand. *Higher Education, 63*(5), 607–620. https://doi.org/10.1007/s10734-011-9461-8

Gawel, A. (2014). Business collaboration with universities as an example of corporate social responsibility-a review of case study collaboration methods. *The Poznan University of Economics Review, 14*(1), 20.

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., & Bell, M. (2005). The micro-determinants of meso-level learning and innovation: Evidence from a Chilean wine cluster. *Research Policy, 34*(1), 47–68. https://doi.org/10.1016/j.respol.2004.10.008

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*(3), 481–510. https://doi.org/10.1086/228311

Granovetter, M. (2005). The impact of social structure on economic outcomes. *Journal of Economic Perspectives, 19*(1), 33–50. https://doi.org/10.1257/0895330053147958

Greene, W. H. (2011). *Econometric analysis* (7th ed.). Pearson Education.

Hemmert, G. A. J., Schons, L. M., Wieseke, J., & Schimmelpfennig, H. (2018). Log-likelihood-based pseudo-R2 in logistic regression: Deriving sample-sensitive benchmarks. *Sociological Methods & Research, 47*(3), 507–531. https://doi.org/10.1177/0049124116638107

Hinzmann, S., Cantner, U., & Graf, H. (2019). The role of geographical proximity for project performance: Evidence from the German leading-edge cluster competition. *The Journal of Technology Transfer, 44*(6), 1744–1783. https://doi.org/10.1007/s10961-017-9600-1

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. https://doi.org/10.1016/j.respol.2012.05.012

Huber, F. (2012). On the role and interrelationship of spatial, social and cognitive proximity: Personal knowledge relationships of R&D workers in the Cambridge information technology cluster. *Regional Studies, 46*(9), 1169–1182. https://doi.org/10.1080/00343404.2011.569539

Hughes, R. A., Heron, J., Sterne, J. A., & Tilling, K. (2019). Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. *International Journal of Epidemiology, 48*(4), 1294–1304. https://doi.org/10.1093/ije/dyz032

Lam, A. (2005). Work roles and careers of R&D scientists in network organizations. *Industrial Relations: A Journal of Economy and Society, 44*(2), 242–275. https://doi.org/10.1111/j.0019-8676.2005.00383.x

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. https://doi.org/10.1080/00343400903401618

Leitgöb, H. (2013). The problem of modeling rare events in ML-based logistic regression - assessing potential remedies via MC simulations.

Leitgöb, H. (2020). Analysis of Rare Events. In: P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, & R. A. Williams (Eds), SAGE Research Methods Foundations.

Long, B. T. (2004). How have college decisions changed over time? An application of the conditional logistic choice model. *Journal of Econometrics, 121*(1), 271–296. https://doi.org/10.1016/j.jeconom.2003.10.004

Maietta, O. W. (2015). Determinants of university–firm R&D collaboration and its impact on innovation: A perspective from a low-tech industry. *Research Policy, 44*(7), 1341–1359. https://doi.org/10.1016/j.respol.2015.03.006

Mansfield, E. (1995). Academic research underlying industrial innovations: Sources, characteristics, and financing. *Review of Economics and Statistics, 77*(1), 55–65. https://doi.org/10.2307/2109992

Mazzucato, M., & Penna, C. (2016). The Brazilian innovation system: a mission-oriented policy proposal. CGEE.

McFadden, D. (1979). Quantitative methods for analysing travel behaviour of individuals: Some recent developments. In D. A. Hensher & P. R. Stopher (Eds.), *Behavioural travel modelling* (pp. 279–318). Croom Helm.

Ministry of Economics. (2015). *RAIS - Annual Social Information Report [Relação Anual de Informações Sociais - confidential dataset]* Ministry of Economics.

Muscio, A. (2010). What drives the university use of technology transfer offices? Evidence from Italy. *The Journal of Technology Transfer, 35*(2), 181–202. https://doi.org/10.1007/s10961-009-9121-7

Muscio, A. (2013). University-industry linkages: What are the determinants of distance in collaborations? *Papers in Regional Science, 92*(4), 715–739. https://doi.org/10.1111/j.1435-5957.2012.00442.x

Muscio, A., & Pozzali, A. (2013). The effects of cognitive distance in university-industry collaborations: Some evidence from Italian universities. *The Journal of Technology Transfer, 38*(4), 486–508. https://doi.org/10.1007/s10961-012-9262-y

Nerad, M., & Evans, B. (2014). Globalization and its impacts on the quality of PhD education: Forces and forms in doctoral education worldwide. Sense Publishers. https://books.google.com.br/books?id=fbP0AwAAQBAJ

Nooteboom, B. (2002). *Trust: Forms, foundations, functions, failures and figures*. Edward Elgar Publishing.

Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., & Van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy, 36*(7), 1016–1034. https://doi.org/10.1016/j.respol.2007.04.003

OECD. (2019). *University-industry collaboration: New evidence and policy options*. OECD Publishing.

Østergaard, C. R. (2009). Knowledge flows through social networks in a cluster: Comparing university and industry links. *Structural Change and Economic Dynamics, 20*(3), 196–210. https://doi.org/10.1016/j.strueco.2008.10.003

Van der Paal, B. (2014). A comparison of different methods for modelling rare events data. *Master of Statistical Data Analysis thesis, Universiteit Gent*.

Partha, D., & David, P. A. (1994). Toward a new economics of science. *Research Policy, 23*(5), 487–521.

Perkmann, M., Salandra, R., Tartari, V., McKelvey, M., & Hughes, A. (2021). Academic engagement: A review of the literature 2011-2019. *Research Policy, 50*(1), 104114.

Petruzzelli, A. M. (2011). The impact of technological relatedness, priorties, and geographical distance on university–industry collaborations: A joint-patent analysis. *Technovation, 31*(7), 309–319. https://doi.org/10.1016/j.technovation.2011.01.008

Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in Regional Science, 86*(3), 423–443. https://doi.org/10.1111/j.1435-5957.2007.00126.x

Ponomariov, B. (2009). Student centrality in university–industry interactions. *Industry and Higher Education, 23*(1), 50–62. https://doi.org/10.5367/000000009787641369

Puhr, R., Heinze, G., Nold, M., Lusa, L., & Geroldinger, A. (2017). Firth's logistic regression with rare events: Accurate effect estimates and predictions? *Statistics in Medicine, 36*(14), 2302–2317.

Roach, M., & Sauermann, H. (2010). A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. *Research Policy, 39*(3), 422–434. https://doi.org/10.1016/j.respol.2010.01.004

Rybnicek, R., & Königsgruber, R. (2019). What makes industry–university collaboration succeed? A systematic review of the literature. *Journal of Business Economics, 89*(2), 221–250. https://doi.org/10.1007/s11573-018-0916-6

Santos, P., Veloso, L., & Urze, P. (2020). Students matter: The role of doctoral students in university–industry collaborations. *Higher Education Research & Development*. https://doi.org/10.1080/07294360.2020.1814702

Skinner, B. T. (2019). Choosing College in the 2000s: An updated analysis using the conditional logistic choice model. *Research on Higher Education, 60*(2), 153–183. https://doi.org/10.1007/s11162-018-9507-1

Tartari, V., Salter, A., & D'Este, P. (2012). Crossing the Rubicon: Exploring the factors that shape academics' perceptions of the barriers to working with industry. *Cambridge Journal of Economics, 36*(3), 655–677.

Teirlinck, P., & Spithoven, A. (2013). Research collaboration and R&D outsourcing: Different R&D personnel requirements in SMEs. *Technovation, 33*(4–5), 142–153. https://doi.org/10.1016/j.technovation.2012.11.005

Ter Wal, A. L., & Boschma, R. A. (2009). Applying social network analysis in economic geography: Framing some key analytic issues. *The Annals of Regional Science, 43*(3), 739–756. https://doi.org/10.1007/s00168-008-0258-3

Thune, T. (2009). Doctoral students on the university–industry interface: A review of the literature. *Higher Education, 58*(5), 637. https://doi.org/10.1007/s10734-009-9214-0

Thune, T. (2010). The training of “triple helix workers”? Doctoral students in university–industry–government collaborations. *Minerva, 48*(4), 463–483. https://doi.org/10.1007/s11024-010-9158-7

UNESCO-UIS. (2015). International Standard Classification of Education - Fields of education and training 2013 (ISCED-F 2013) – Detailed field descriptions. UNESCO Institute for Statistics. Doi: https://doi.org/10.15220/978-92-9189-179-5-en

Vinding, A. L. (2004). Interaction between firms and knowledge institutions. In R. A. Burgelman & H. Chesbrough (Eds.), *Product inovation, interactive learning and economic performance.* Emerald Group Publishing.