The role of academic relations of former graduate students in university-firm collaboration: evidence from Brazil

Daniel Gama e Colombo[*]

Renato de Castro Garcia[†]

Abstract

This paper investigates the contribution of the personal ties of former Master and Ph.D. students to university-firm collaboration. Using the proximity framework developed by Boschma (2005) and the underlying assumptions of social proximity (trust, commitment, common language and common culture), we argue that the academic relations these former students developed during graduate education can reduce the social distance between universities and firms, thus favoring collaborative research. Based on this argument, we present two hypotheses to explain how hiring a former graduate student is associated with the collaboration decisions of private organizations. These hypotheses are tested with a new empirical strategy, using a novel and comprehensive dataset on university-industry linkages in Brazil, and modelling the private organization's decision in two steps, i.e., the choice of a partner and the decision to collaborate. We find that, if a research group is hosted by a university in which one or more employees of a private organization attended graduate education, the employer organization is more likely to choose such group to partner (relative odds around 2.5 times higher) and to engage in collaboration (odds ratio more than 4 times higher). We also find that the magnitude of this association varies substantially per broad field of education, supporting the proposition that scientific disciplines work as 'moderators' of the social dimension of proximity. These results are the main contributions of the paper to the understanding of university-firm collaboration, and they suggest new approaches for policy support to these partnerships, using academic relations as a lever to new collaborative projects.

Keywords: conditional logit; graduate education; social proximity; university-firm collaborations.

JEL classification: I23; O30; O31

[*] Visiting researcher at the Institute of Economics of the University of Campinas. E-mail: dcolombo@alumni.usp.br.

[†] Professor at the Institute of Economics of the University of Campinas – rcgarcia@unicamp.br.

O papel das relações acadêmicas de ex-estudantes de pós-graduação na colaboração universidade-empresa: evidências do caso brasileiro

Daniel Gama e Colombo[*]

Renato de Castro Garcia[†]

Resumo

Este artigo analisa a contribuição dos laços pessoais de ex-estudantes de mestrado e doutorado para a colaboração universidade-empresa. Com base no arcabouço de proximidade desenvolvido por Boschma (2005) e nas premissas do conceito de proximidade social (confiança, compromisso, linguagem comum e cultura comum), propõe-se que as relações acadêmicas que esses ex-alunos desenvolveram durante a pós-graduação podem reduzir a distância social entre universidades e empresas, favorecendo a pesquisa colaborativa. À luz desse argumento, são apresentadas duas hipóteses para explicar como a contratação de um ex-aluno de pós-graduação está associada à decisão de colaborar de uma organização privada. Essas hipóteses são testadas a partir de uma nova estratégia empírica, utilizando uma nova e abrangente base de dados sobre as parcerias entre universidades e empresas no Brasil, e modelando a decisão da empresa em duas etapas, quais sejam, a escolha do parceiro e a decisão de colaborar. Os resultados indicam que, se um grupo de pesquisa pertencer a uma universidade na qual um ou mais empregados de uma organização privada tenham frequentado a pós-graduação, há maior verossimilhança de que essa organização escolha esse grupo de pesquisa como parceiro (razão de chances cerca de 2,5 vezes maior) e decida colaborar (razão de chances mais de 4 vezes maior). Além disso, a magnitude encontrada dessa associação varia de acordo com a 'grande área de conhecimento' em questão, indicando que a área de conhecimento pode constituir um moderador da proximidade social. Esses resultados são as principais contribuições do artigo para a compreensão da colaboração universidade-empresa, e sugerem novas abordagens para políticas públicas para apoiar essas parcerias, que utilizem as relações acadêmicas como alavancas para novos projetos colaborativos.

Palavras-chave: colaboração universidade-empresa; logit condicional; pós-graduação; proximidade social.

Classificação JEL: I23; O30; O31

[*] Pesquisador-colaborador do Instituto de Economia da Universidade de Campinas. E-mail: dcolombo@alumni.usp.br.

[†] Professor do Instituto de Economia da Universidade de Campinas. E-mail: rcgarcia@unicamp.br.

1 Introduction

Graduate education[1] constitutes the most advanced level of academic training, and it provides students not only with deep knowledge and analytical skills, but also with opportunities to interact with and develop relations with their peers, professors, and other members of the academic community. Such relations constitute important connections for the alumni, and they can foster new projects and partnerships after they obtained their degree. For this reason these connections are encouraged by firms and universities alike (Sauermann & Stephan, 2010). Although the importance of such ties for collaborative research are known and acknowledged, their empirical confirmation and measurement remains a gap in the literature. This paper aims at contributing to the understanding of university-firm collaboration for innovation projects, discussing how academic relations can help and are linked to the collaboration decisions of private organizations.

Graduate education has been growing steadily in the last decades (Nerad & Evans, 2014). There is substantial empirical evidence of its benefits to individuals in terms of human capital accumulation, employment opportunities (especially in non-routine jobs), higher income and wages, and access to corporate 'fast-track' careers (Kingston & Clawson, 1985; Lindley & Machin, 2016; Mertens & Röbken, 2013; Zhang, 2005). However, the evidence of its contribution to economic performance and to society as a whole is far more limited,[2] remaining a gap in the literature (Halse & Mowbray, 2011; Raddon & Sung, 2009), requiring requires further investigations aimed at measuring the positive outcomes of graduate education, and explaining the channels through they take place.

One of the potential contributions that needs to be examined more carefully is how former students of graduate programs help fostering university-firm collaborations, an important driver of firms' innovative efficiency (Cosh, Fu, & Hughes, 2005), resource allocation and industry competitiveness (Cunningham & Gök, 2012). Firms hiring such professionals after they obtained their degree not only increase their ability to evaluate, assimilate and exploit external knowledge (i.e., their absorptive capacity - Cohen & Levinthal, 1990), but they also get access to the network of these 'linked scientists' within

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

[2] Some of the benefits of graduate education discussed in the literature are the generation of basic knowledge and productivity boost (Casey, 2009), higher level skills that are key drivers to innovation, entrepreneurship (Leitch, 2006), advanced knowledge and capabilities, and social and cultural experience (Smith et al., 2010).(Hossler, Braxton, & Coopersmith, 1989)

the academic community (Lam, 2005), making them a key component of university-industry collaborations (Ponomariov, 2009).

Although this subject has been discussed in previous studies (Freitas, Marques, & e Silva, 2013), the contribution of personal ties of former graduate students to these partnerships under the assumptions of the proximity framework (Boschma, 2005) remains unclear. Previous analyses within this literature have acknowledged the relevance of personal relationships for collaboration, but they have not properly addressed the importance of graduate education for social proximity.[3] Drejer and Østergaard (2017) and Østergaard (2009) used the connections of firms' employees within their universities at the undergraduate level ('employee-driven relations'), yet relationships developed during the Master's or Ph.D. level are not considered. As a result, it is still necessary to test the contribution of these relations by quantitative model-based analysis, that controls for different dimensions of proximity (Boschma, 2005), along with individual features of both partners.

This paper aims at filling this gap in the literature by investigating how academic relations of former graduate students can reduce the social distance between universities and firms, thus favoring collaborative research. We explain how and test whether hiring a former graduate student is associated with the likelihood of a private organization collaborating with a research group belonging to the employee's graduate university. Our contribution to the existing literature is twofold. First, we incorporate such academic relations as drivers of collaborative research, based on the underlying assumptions of social proximity, such as trust, commitment, common language and common culture. Second, we test and measure the importance of such relations with a new empirical strategy, using a novel and comprehensive dataset on university-industry linkages in Brazil, and modelling firm's decision in two steps, as choice of partner and decision to collaborate.

A difficulty in assessing the importance of proximity factors is that it may vary depending on the knowledge field of the research group. Rybnicek and Königsgruber (2019) argued that scientific disciplines can work as 'potential moderators' of their effects, because of distinct conventions, methods, and openness of different fields to the needs of industry. However, the authors maintain that this remains a large gap in the

[3] Existing studies have used other variables to measure social proximity, such as previous collaborations (Broekel, 2015; Cassi & Plunket, 2014; Hong & Su, 2013; Petruzzelli, 2011), and work experience of top managers (Broekel & Hartog, 2013).

literature, as most empirical studies limit the analysis to a specific field or sector[4], or simply do not address this issue, which may bias the estimates or limit the generalization of findings. A third contribution of this analysis to the literature is to shed a light on this point by estimating the parameters of the model for each 'broad field of education and training' (UNESCO-UIS, 2015) separately, along with an estimate for all fields.

The empirical analysis presented herein is based on data for the Brazilian case. The landscape of innovation and university-firm collaboration in the country has been extensively discussed and described in previous studies. Brazil has a low but heterogeneous innovative base, with a small group of excellence research centers.[5] The public sector has been crucial for the formation of the scientific and technological base, but public funding has been drastically reduced in the last decade (Cassiolato, 2015). Despite its weaknesses (low propensity to innovate, fragmentation and lack of long-term agenda), the country presents elements of a developed system of innovation, along with strategic natural assets and a strong domestic market (Mazzucato & Penna, 2016). University-firm collaborations for innovation projects are highly concentrated in the southern part of the country (Garcia, Araujo, Mascarini, Gomes Santos, & Costa, 2015), with comparatively high number of projects in engineering and agrarian sciences (Suzigan, Albuquerque, Garcia, & Rapini, 2009), and public support focused in mature industries (Freitas et al., 2013).

The remainder of this paper is structured as follows: the second section following this introduction briefly discusses the main developments and the state-of-the-art of the economic geography literature; the third part presents the main arguments and research hypotheses, that are tested according to the empirical analysis described in the fourth section; the fifth part 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.

[4] Broekel and Hartog (2013) considered only the Dutch aviation industry, Autant-Bernard, Billand, Frachisse, and Massard (2007) focused on micro and nanotechnologies,

[5] Such as the Brazilian Agricultural Research Corporation – EMBRAPA, and the PETROBRAS' Research and Development Center – CENPES.

2 Dimensions of Proximity: Literature Review

The evolution of knowledge networks has received a growing attention in economic geography (Broekel, 2015; Ter Wal & Boschma, 2009). Recent studies are not only trying to explain the underlying dynamics of network evolution, but also to understand whether geographical proximity still plays an important role to the establishment of collaboration ties among different partners.

Geographical proximity facilitates linkages among partners, due to the existence of mechanisms such as frequent interactions and face-to-face contacts. Firms often prefer to collaborate with close universities, since their search processes for academic partners often point to closer universities, which allows them to reduce the costs of the partnership, and facilitates closer interactions with their academic partners (D'Este, Guy, & Iammarino, 2013). Another reason for this preference is that university's knowledge spillovers are geographically bounded, so firms can capture more results of research and development projects under development (Audretsch & Feldman, 1996). Firms also tend to favour local universities as academic partners both because they know local researchers' projects and activities, as a result of social ties (Drejer & Østergaard, 2017), and because they can reduce the costs of academic collaboration (Muscio, 2013). These reasons affect firms' search processes and increase the impact of geographic proximity. Local universities are frequently a vital and very visible agent of a region's identity, building networks, qualifying students, and being part of the agents' local ties.

But there are other dimensions of proximity that can foster interactive learning among partners, that can be summarised in: cognitive, social, organizational, and institutional proximity. Empirical evidence confirms the (simultaneous) relevance of all proximity types as significant drivers of network evolution (Balland, 2012; Cassi & Plunket, 2014).

Cognitive proximity is related to the level of overlap in two actors' knowledge bases (Nooteboom, Van Haverbeke, Duysters, Gilsing, & Van den Oord, 2007). Actors need to have a complementary absorptive capacity to identify, interpret and exploit the knowledge of their partners, but different cognitive bases and different absorptive capacities are necessary to explore new knowledge (Cohen & Levinthal, 1990). Since agents have similar knowledge bases, there are similarities in the way the world is perceived, interpreted, understood, and evaluated by them. However, if the overlap is too strong, interaction will rarely result in new combinations, due to the lack of novelty (Nooteboom, 2000). In fact, we can find what is known as the "proximity paradox": a high degree of cognitive

proximity is a driving factor for interactive learning, but it does not necessarily improve innovative performance, and may even harm it (Boschma & Frenken, 2010). This assumption suggests that there is an optimal level of cognitive proximity to trigger innovation (Cassi & Plunket, 2014).

Social proximity refers to the strength of interpersonal linkages, or to what extent 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). Of particular interest is the role of trust, which is likely to be positively influenced by social proximity, and it is frequently said to foster knowledge exchange (Broekel, 2015; Nooteboom, 2002). The main argument is that strong, trust-based ties facilitate knowledge sharing and interactive learning (Gertler, 2003; Huber, 2012). Other arguments that support the importance of social proximity for collaboration are: sharing of common language, which may be crucial for an effective communication to develop new ideas and technologies, and for reducing information incompleteness or asymmetry (Baba, Yarime, & Shichijo, 2010; Gawel, 2014; Lin, Geng, & Whinston, 2005; Vanhaverbeke, 2006); common culture (Barnes, Pashby, & Gibbons, 2002) and commitment (Attia, 2015). Empirical studies have shown that social proximity increases the likelihood of linkages among actors, that is, when two individuals have a partner in common, they are more likely to end-up forming a collaboration (Cassi & Plunket, 2014; Huber, 2012). Closer connection facilitates communication and interactive learning, since it gives better access to knowledge, and therefore, it may also increase innovation. Said differently, social proximity provides individuals advantages over more distant agents; distant sources of knowledge certainly provide diversity and new ideas, but incur costs to combine and manage different sources of knowledge (Cassi & Plunket, 2014).

The other two dimensions of proximity addressed by the literature are institutional and organizational. Although the concepts are related, the first indicates the degree to which two institutions are subject to the same institutional framework, background, and systems of rewards and values (Broekel, 2015; Ponds, Van Oort, & Frenken, 2007). Organizational proximity, on the other hand, 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).

The literature on proximity dimensions reviewed in this section provides a complex framework that allows one to consider how different factors are associated with

scientific collaboration within a country or region. However, up to this date, such literature has not considered how personal relations of former graduate students can influence such partnerships. In the next section, we develop a set of hypotheses on the subject using this framework, and, particular, the ideas and arguments related to social proximity.

3 Research hypotheses

The hypotheses developed in this section are based on the idea that former graduate students have "academic relations" within the university they studied (at graduate level). The idea of academic relations refers to the social ties of Master's or Ph.D. students with their advisors, professors, peers and other members of their universities, developed during the course of graduate training through classes, seminars, meetings and other channels of social and academic interaction. Students who leave academia to work for other organizations still nurture such relations through activities as publications, professional societies and attendance to academic congresses and professional meetings (Roach & Sauermann, 2010; Sauermann & Stephan, 2010).

Such academic relations are valuable to the firm not only because they are an important source of knowledge and absorptive capacity (Roach & Sauermann, 2010), but also because they act as catalysts for partnerships between institutions. From the perspective of the university, these academic relations are important connectors of the academia to the 'outside world' (Balconi & Laboranti, 2006), as former pupils working at the private sector are preferential links for professors, both because of personal relations and cognitive proximity. With personal connections in (and valued by) both the academic and private sectors, these former students play a crucial role in university-firm collaboration (Ponomariov, 2009), as 'linked scientists' that constitute 'knowledge network nodes' between the private sector and universities (Lam, 2005), and, in some cases, taking the initiative of new collaborative projects (Freitas et al., 2013).

In our analysis, these relations are taken as the social dimension of proximity for university-firm collaboration. The literature suggests several potential reasons to explain how such academic relations may be associated with university-firm R&D collaboration: first, the 'relational capital' built by former graduate students signals trust and respect, smoothing negotiations and constituting a driver of partnerships between private organizations and the scientific community (Attia, 2015; Canhoto, Quinton, Jackson, &

Dibb, 2016; Teirlinck & Spithoven, 2013); second, these employees are likely to share a common language with their university peers and professors (Baba et al., 2010; Gawel, 2014; Lin et al., 2005; Vanhaverbeke, 2006); similarly, Masters and Ph.D.'s working at the industry can reduce the cultural gap with universities, balancing different priorities, goals and timing of the organizations involved (Barnes et al., 2002); finally, employees with graduate training may strengthen the commitment of firms for collaboration, increasing the willingness to allocate effort and resources to these projects (Attia, 2015).

We assume that an organization employing a former graduate student from a university accesses his or her academic relations, thus affecting its incentives, costs and expected returns of collaboration. Such employee would work as a link, increasing the social proximity between the employer organization and the university. Based on such theoretical arguments, we present the following hypotheses on how employing a former graduate student can be associated with the likelihood of collaboration between universities and private organizations:

Hypothesis 1 - private organizations are more likely to choose a research group to collaborate with if one or more of its employees have attended graduate education (Master's or Ph.D.) at the group's host university.

Empirical studies on university-firm collaborations investigated the factors determining the choice of a partner or the formation of a dyad between two institutions (Broström, 2010; Carayol, 2003; D'Este et al., 2013; De Fuentes & Dutrénit, 2012; Garcia, Araujo, Mascarini, Gomes Dos Santos, & Costa, 2018; Hong & Su, 2013; Isabel Maria, Rossi, & Geuna, 2014; Petruzzelli, 2011). In general, these analyses considered only organizations that actually engaged in collaboration (based on patent citations, surveys or administrative data), in order to identify and test the features and factors explaining the formation of a link between two institutions.

We expect the social proximity induced by academic relations to reduce the perceived risks of a potential partner related to its innovation capabilities, opportunistic behavior or lack of engagement. The former student in the workforce of the private organization can signal trust and commitment to the university researchers (Attia, 2015; Canhoto et al., 2016; Teirlinck & Spithoven, 2013), both because of previous interactions, long-term relations and of prospects of future collaborations. Such employee is also in a better position to forecast the potential results of the collaboration, as he or she is 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, facilitating the negotiations for knowledge transfer, allocation of resources, and division of project results. Consequently, these academic relations add value to the expected return of the projects, making the collaboration with research groups from these universities more promissing, and therefore more likely, as suggested by *'Hypothesis 1'*.

Hypothesis 2 - private organizations 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.

A much less studied topic in this literature is the investigation of factors associated with the firm's decision to engage (or not) in collaboration with any university. Such analysis requires a richer sample that includes both organizations that coolaborated and those that did not. Yet, this constitutes a promissing research agenda to be explored, as the few existing studies have found evidence of the correlation of distinct factors with such decision, including geographical proximity (Drejer & Østergaard, 2017), knowledge intensity (Hanel & St-Pierre, 2006), degree programmes (Maietta, 2015), innovation capabilities (De Fuentes & Dutrénit, 2012), 'open' search strategies (Laursen & Salter, 2004), absolute size and degree of openness (Fontana, Geuna, & Matt, 2006).

Once again, we expect academic relations of former graduate students to be associated with this decision. As a private organization hires such employees, it improves the social proximity by reducing the cultural and language gap and potential information assymetries with research groups from the university where they studied (Baba et al., 2010; Barnes et al., 2002), besides improving the perception of trust and commitment. These employees are also likely to share the same or a similar technological paradigm with researchers from their universities (i.e., a higher cognitive proximity), facilitating the communication and identification of problems and technological solutions that are both commercially applicable and academically promissing, so that it presents benefits to both parties of the transaction. For these reasons, private organizations interested in or seeking collaboration with research groups in universities where their employees studied (at the Master's or Ph.D. level) are more likely to pursue such strategy for technological development, as suggested by *'Hypothesis 2'*.

The hypotheses presented in this section suggest a reasoning for arguing that academic relations are relevant and therefore should be associated with the choices and decisions of firms for scientific collaboration with universities, along with other proximity factors and features of the collaborating parties. Such propositions are tested empirically, as described in the next section.

4 Empirical Strategy and Data

The objective of the empirical analysis presented in this section is to test the above research hypotheses and to investigate the role of former students of graduate programs (Masters and Ph.D.'s) to help building the bridge for collaborations between private organizations and universities' research groups. The main goal is to test whether hiring a former graduate student can help to predict whether a private organization will collaborate a with a research group belonging to the employee's university. Considering the available data (described in item 4.2), private organizations are interpreted herein as including commercial firms, nonprofit private organizations, public companies and public nonprofit organizations under private law.

4.1 The empirical model

In order to test the abovementioned hypotheses, we use a two-step model to describe the private organization's decision-making process for collaborating with a research group. The empirical literature on the subject usually considers only one of these stages (De Fuentes & Dutrénit, 2012), applying suitable estimators.[6] But sequential decisions are not disconnected, and it is reasonable to suspect that the choice made at earlier stages affects the next ones. Although uncomon, multiple-stage models are not unprecedented in this literature: De Fuentes and Dutrénit (2012) devised a three-stage model for the drivers, channels and benefits of collaboration; and Laursen, Reichstein, and Salter (2011) distinguished between the decision to collaborate and the choice of collaborating with a local university.

[6] The most common estimators used in this literature include binomial or multinomial logit and logistic regression (Autant-Bernard et al., 2007; Cassi & Plunket, 2014; D'Este et al., 2013; Drejer & Østergaard, 2014; Isabel Maria et al., 2014; Laursen & Salter, 2004), probit (Broström, 2010; Maietta, 2015), panel data estimators (Colombelli, Krafft, & Quatraro, 2014), and ordinary least squares (Garcia et al., 2018).

The two-step decision model applied herein was originally developed to explain higher education choice, and it was presented in detail by Long (2004) and Skinner (2019). We adapted this framework to the context of university-firm collaboration, an original approach not considered in previous studies. The model links both stages of the decision-making process, providing a more complex description of the drivers and factors influencing collaboration. The main advantage is the possibility to measure and test whether social proximity can predict the decision to engage in collaboration (*Hypothesis 2)*, by inputting the results of the first stage into the second, as described below. It also has the advantage of overcoming the computational and statistical difficulties of working with an enormous dataset that presents several possible matches for each private organization (Cassi & Plunket, 2014; Sorenson, Rivkin, & Fleming, 2006), without resorting to an arbitrary number of controls for each dyad,[7] which can bias and jeopardize the consistency of the estimates.

The decision to collaborate with a research group for an innovation project is divided in two stages, and each one tests one of the above research hypotheses. The private organization first considers all research groups available for collaboration, identifying the one with highest expected net result. In the second stage, it compares such result with the option of not collaborating (i.e., not developing R&D or developing it internally), deciding the best course of action. In light of the structure and updating procedures of the databases used in the analysis, we assume that collaboration decisions at year '*t'* are based on attributes of the research groups and host universities at the same period, and on features of the private organizations and employees' educational attainment observed at period '*t -1'*.

We expect that the dimensions of proximity described by Boschma (2005) and Boschma and Frenken (2010) are associated with the decision to collaborate in both stages. For this reason, we introduce variables representing the geographical, social and institutional dimensions (listed in item 4.2).[8] To control for cognitive proximity, we follow Ponds et al. (2007) and assume that the search for a partner at the first stage is

[7] Sorenson et al. (2006) and Hong and Su (2013) use a 'case-control design', determining a specififc number of non-realized collaborations (controls) for each actual realized collaboration.

[8] Organizational proximity (Balland, 2012) is not considered in this study because universities and private organizations seldom rarely belong to the same corporate group in Brazil, thus not being a relevant factor to be considered in the model.

limited to a specific knowledge field,[9] so that all potential research groups available for collaboration present a (similar) small cognitive distance.[10]

4.1.1. The first stage: choice of partner

At the first stage, private organizations search among all potential research groups (i.e., those belonging to a specific knowledge field), considering the expected costs, returns and risks of each choice. The objective of this stage is to identify the most rewarding collaboration. To make this decision, the organization takes into consideration features of the research group and its host university, along with the proximity factors mentioned previously. In particular, we expect that the academic relations of former graduate students add value to the collaborations with research groups from their universities. As a result, employing these workers would be positively associated with the relative odds of the private organization choosing such groups, according to *Hypothesis 1*.

As proposed by Long (2004) and Skinner (2019), the decision at this stage is modelled as a probabilistic equation. The probability (P_{ij}) that a randomly drawn private organization i will choose a research group j ($choice_i = j$) as the one with highest net expected result is (Greene, 2011):

$$P_{ij} = \text{Prob}\left(choice_i = j | X_{ij}\right) = \frac{exp\left(\beta' x_{ij}\right)}{\sum_{k=1}^{J} \exp\left(\beta' x_{ik}\right)} \quad (1)$$

Where X_{ij} is a vector of explanatory variables associated with the probability of a private organization choosing a particular research group, presented in Table 1 as '*attributes of the research group and host university'* and '*proximity factors'*; 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; McFadden, 1973), suitable for cases in which the decision-maker is faced with a great number of choices, as it exploits the variation of attributes and

[9] The 'knowledge field' refers to the 'main field of education' reported by each research group to the '2016 Census of Research Groups of the National Council for Scientific and Technological Development' (CNPq, 2016).

[10] This strategy has the additional advantage of limiting the size of the dataset necessary for estimation, solving the computational problem reported by Hong and Su (2013) and Sorenson et al. (2006). Each private organization is paired only with the research groups of the same knowledge field as the one the organization actually collaborated with.

interaction terms (Long, 2004). This model also has the advantage of controlling for individual attributes of the private organizations, as they are differenced out of the equation.

For this estimation, we use data on collaborations that actually took place (as described in item 4.2). It is assumed that all private organizations have decided for the collaborations with highest expected return. The initial dataset is expanded to cover all potential choices of each private organization that actually collaborated with a research group, i.e., all possible dyads of one of these organizations and an existing research group within the relevant field of the actual partner observed in the original dataset. The resulting expanded database includes a dummy that informs the realized ties (the original choice or collaboration), which constitute the outcome variable of the probabilistic model.

Consistency of the estimates of the conditional logit model depends on the strong assumption of independence of irrelevant alternatives (IIA), that basically requires that the odds ratio between two alternatives is not affected by the inclusion (or exclusion) of other alternatives (Train, 2003). This assumption represents the major shortcoming of the estimator, especially in the cases of close substitutes (McFadden, 1973; Train, 2003). Unfortunately, existing tests of the IIA assumption have been found to report inconsistent results for applied research (Cheng & Long, 2007). Yet, based on arguments presented by Skinner (2019), we understand that we have good grounds to maintain that the IIA assumption should not pose a threat of bias in this case: first, as all potential choices of collaboration for each private organization are considered ('completeness of the choice set'), a potential bias caused by omission of a relevant option is unlikely; second, collaboration with each research group constitutes an independent and unrelated transaction, so that the odds of choosing between two or more groups should not be changed by the potential existence of an additional alternative; and third, research groups have very different and unique features (e.g., number and disciplinary specialty of researchers, main topics of research, funding, research infrastructure), so that they represent very distinct choices and can be hardly considered 'close substitutes', minimizing the IIA assumption problem, as suggested by Train (2003).

4.1.2. The second stage: decision to collaborate

At the second stage, the private organization considers the expected costs and earnings of its best choice of collaboration (selected at the first stage), comparing it with the alternative of not engaging in collaboration. In this decision, the organization weights

its own features and absorptive capacity (Cohen & Levinthal, 1990), that is considered a necessary requirement to benefit from knowledge obtained through collaboration (Autant-Bernard et al., 2007). For analytical purposes, this stage encompasses not only the private organization's decision, but also other Williamsonian transaction costs (Williamson, 1985), such as project negotiation, intellectual property issues, and any other obstacles that need to be overcome for the collaboration to occur. Again, we expect proximity factors and academic relations of former graduate students to be associated with the expected return of the collaboration, thus constituting predictors of this decision, as suggested in *Hypothesis 2*.

To investigate these factors and estimate their association with the decision at this stage, we must have in the dataset private organizations that did not collaborate. However, in these cases the research group chosen as the best option (necessary to calculate the proximity variables) is not known. To overcome this problem, we pair each private organization with an estimated 'most likely partner', as follows: first, a knowledge field is selected for each private organization that did not collaborate, using the most frequent choice of organizations of the same sector (the mode of each sector)[11]; then, we expand the dataset to cover all potential collaboration choices for each private organization (meaning all research groups within the respective knowledge field, thus controlling for cognitive distance); third, we apply the β parameters estimated at the first stage (for each broad field) to all dyads in the dataset; finally, the research group with highest probability (P_{ij})[12] is selected as the 'most likely partner' and paired with the respective private organization, excluding all other potential choices. For private organizations that actually collaborated, the most likely partner is not necessarily the one in the original dataset.

The response variable of the second stage is a dummy that represents the decision of the organization i to engage in collaboration ($collaboration_i = 1$) or not ($collaboration_i = 0$). The decision is again modelled as a probabilistic function, where the independent variables are the ones presented in Table 1 as '*features of the private organization'* (including its absorptive capacity, as discussed previously), '*attributes of the research group and host university*', considering the estimated 'most likely partner', and *'proximity factors'*. Following previous studies that investigated this decision (Drejer & Østergaard, 2017; Laursen & Salter, 2004) and considering the binary nature of the

[11] National Classification of Economic Activities (CNAE), 2-digit level.
[12] No minimum probability cutoff was used for this estimation.

dependent variable, we apply a standard logistic regression (Greene, 2011) to estimate the associations of different factors with the likelihood of the outcome.

4.2 Data, sample design and descriptive statistics

To estimate the empirical model and test the research hypotheses, we present a novel and rich dataset that comprises microdata on a large number of private organizations in Brazil, their collaboration with academic research groups, and employment of former graduate students. This database constitutes an important contribution of this research, as many empirical analyses on university-firm collaboration are based on small samples or are limited to collaborations established by few universities (Broström, 2010; D'Este et al., 2013; Drejer & Østergaard, 2017; Isabel Maria et al., 2014). Also, the large number of organizations in our dataset makes it representative of such partnerships in Brazil, strengthening the evidence presented for the hypotheses.

To generate the dataset, we merged information from three databases. The first one is the '2016 Census of Research Groups of the National Council for Scientific and Technological Development' (CNPq, 2016), that provides detailed microdata on all research groups in Brazil active in 2016, including institutional affiliation (host university) and collaboration projects under development. Using the National Register of Legal Entities (C.N.P.J.) of private organizations that partnered with these groups, we merged such database with the 2015 version of the 'Annual Social Information Report – RAIS' (Ministry of Economics, 2015), which contains information on employment contracts of all legal entities in Brazil. Finally, we used the Individual Taxpayers' Register (C.P.F.) to identify the employees of these organizations that attended graduate education in a Brazilian university from 1996 to 2015, using the database of graduate students of the Coordination of Improvement of Higher Education Personnel (CAPES, 2017), the government office responsible for regulating graduate education in Brazil.

The sample used in the study consisted of: (a) collaborations reported by research groups within Brazilian universities with private organizations (as defined previously; collaborations with public institutions are not included) in 2016; and (b) private organizations with at least one active employee with a higher education degree at the end of 2015. The employees used to assess social proximity are all students who enrolled in a Master's or Ph.D. program from 1996 to 2015 (according to CAPES' 2017 database, and regardless of whether they obtained a degree or not) and were employed by a private organization in 2015. The study follows a 'complete case analysis' approach to deal with

the missing data problem (Hughes, Heron, Sterne, & Tilling, 2019; Seaman & White, 2013), meaning that the sample is limited to units for which full information is available.

The following research choices were also necessary to ensure the feasibility of the analysis: (a) first, the units of observation are individual business units of private organizations;[13] (b) collaborations between research groups within the same university are not included; and (c) the main fields of education reported by research groups were used to classify them in 'broad fields of education and training', following the classification of broad, narrow and detailed fields presented in UNESCO-UIS (2015). The limitations imposed by these choices on the strength of the evidence and generalization of findings are discussed below.

Using the mentioned procedures and choices, the resulting dataset comprises 8,062 collaborations by 3,225 private organizations and 4,738 research groups in 377 universities, along with several other private organizations (more than two million) that did not collaborate with a university in the relevant period.

The dependent variable of the first stage is a dummy that indicates the research groups with which private organizations chose to collaborate in the relevant year '*t*' (2016). The main parameters of interest are the coefficients of proximity factors of each dyad. The social proximity is represented by a dummy indicating whether the private organization employed one or more former graduate students from the same university that hosts the research group in *'t – 1'*, regardless of the program or knowledge field attended by such employee. The geographical proximity is considered by its inverse, i.e., the distance (in 100 kilometers) between the cities of each part of the dyad (Autant-Bernard et al., 2007; D'Este et al., 2013; Laursen et al., 2011; Petruzzelli, 2011). Following the rationale of institutional proximity used in the literature (Balland, Boschma, & Frenken, 2015; Broekel, 2015; Ponds et al., 2007), we measure this dimension as a dummy indicating whether the host university is incorporated under private law, so that it is subject to a similar set of rules and legal statutes as the private organization. The other control variables used to estimate the parameters at this stage are attributes of the research group and host university (observed at period *'t'* – 2016), and state dummies to indicate their location. Features of the private organization and

[13] The same unit of observation was considered in D'Este et al. (2013) and Laursen and Salter (2004). This choice was necessary to include the geographical distance between partners in the model, as one organization or firm can have more than one business unit located in different places.

absorptive capacity are not included, as they are differenced out of the equation of the conditional logit model (Long, 2004).

At the second stage, the dependent variable is a dummy informing whether the private organization engaged in collaboration or not, while the explanatory variables are the abovementioned proximity factors (considering the 'most likely partner' of each private organization), the other independent variables used at the first stage[14], and features of the private organizations observed in *'t – 1'*, including sector dummies[15]. We proxy the absorptive capacity of private organizations using both the share of employees with higher education (following Garcia et al., 2018 and Drejer and Østergaard, 2017) and with a graduate degree, providing a measure of prior knowledge within the organization (Lane, Koka, & Pathak, 2006). The estimated coefficients for graduate degree personnel provide a second set of results for discussing the role of graduate education in scientific collaboration.

All variables used in the study are listed in Table 1. For comparison purposes, descriptive statistics are presented for the entire sample and for private organizations that collaborated only. Table 2 presents the distribution of collaborations in the dataset, with the number of collaborations, private organizations and research groups per broad field of education. Our data indicates that 'Natural Sciences, Mathematics and Statistics' has become an important field for collaborations in Brazil, along with engineering and agrarian sciences, already mentioned in previous papers (Suzigan et al., 2009).

5 Results and discussion

Tables 3 and 4 present the estimated parameters for the first and second stages of the model in odds ratios, that informs 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 (odds are higher if the estimated odds ratio is greater than one, and they are lower if estimates are below unity). Column 1 in both tables presents the estimates for the entire sample, while the others limit the sample to a particular broad field of education and training. Statistical significance of the parameters is assessed at a 0.05 significance threshold (a 95% confidence interval).

[14] At the second stage, state dummies indicate the location of the private organization (and not the research group).
[15] CNAE, 2-digit level.

Table 1. Variables used in the empirical analysis and descriptive statistics (units of observation: business units of private organizations)

Variables	All private organizations	Only private organizations that collaborated
	Mean (Std. Dev.)	Mean (Std. Dev.)
Collaboration with a research group (dummy)	0.001 (0.04)	1 (0)
Number of collaborations		2.51 (10.42)
Features of the private organization		
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: share of employees with an undergraduate degree	0.09 (0.22)	0.41 (0.32)
Absorptive capacity: share of employees with a graduate degree[a]	0.001 (0.02)	0.04 (0.1)
Attributes of the research group and host university		
Host university incorporated as a commercial enterprise (dummy)		0.08 (0.28)
Age of the research group		14.39 (10.57)
Number of researchers of the research group[b]		13.03 (12.64)
Number of private partners of the research group		7.33 (12.39)
Proximity factors		
Geographical distance (per 100 km)[c]		3.17 (5.75)
Institutional (private host university)		0.24 (0.43)
Social (dummy for employment of former graduate student from the host university)		0.23 (0.42)
No. of obs.	2,247,423	3,225

[a] Master's or Ph.D. degree. [b] Not considered students, external members and technical staff. [c] Distance between municipalities.
Source: CAPES (2017), CNPq (2016), and Ministry of Economics (2015).

Table 2. Distribution of collaborations, private organizations (that collaborated) and research groups per broad field of education and training

Broad field of education and training	Collaborations		Private organizations		Research groups	
	Freq.	%	Freq.	%	Freq.	%
Education	407	5.1	131	4.1	292	6.2
Arts and Humanities	246	3.1	87	2.7	185	3.9
Social Sciences, Journalism and Information	482	6	171	5.3	340	7.2
Business, Administration and Law	409	5.1	154	4.8	281	5.9
Natural Sciences, Mathematics and Statistics	1,551	19.2	549	17	1,012	21.4
Information and Communication Technologies	342	4.2	150	4.7	195	4.1
Engineering, Manufacturing and Construction	2,428	30.1	1,090	33.8	1,078	22.8
Agriculture, Forestry, Fisheries and Veterinary	1,360	16.9	562	17.4	747	15.8
Health and Welfare	823	10.2	325	10.1	597	12.6
Services	14	0.2	6	0.2	11	0.2
Total	8,062	100	3,225	100	4,738	100

Source: CAPES (2017), CNPq (2016), and Ministry of Economics (2015).

Table 3. Estimated parameters of the first stage – choice of partner. Conditional Logit Model (with robust variance-covariance matrix). Dependent variable: *choice*$_i$ *(*dummy for research group and host university that collaborated with each private organization).

	Broad Fields of Education and Training[a]									
Independent variables	(1) All fields	(2) Education	(3) Arts and Humanities	(4) Social Sciences, Journalism and Information	(5) Business, Administration and Law	(6) Natural Sciences, Mathematics and Statistics	(7) Information and Communication Technologies	(8) Engineering, Manufacturing and Construction	(9) Agriculture, Forestry, Fisheries, and Veterinary	(10) Health and Welfare
Proximity factors										
Geographical distance (100 km)	0.789***	0.779***	0.742***	0.791***	0.775***	0.779***	0.832***	0.795***	0.769***	0.774***
	(0.004)	(0.020)	(0.034)	(0.017)	(0.021)	(0.010)	(0.020)	(0.009)	(0.009)	(0.016)
Institutional (private host university)	1.147***	1.348**	1.398	1.399**	1.106	1.289***	1.239	1.072	1.039	0.970
	(0.041)	(0.181)	(0.296)	(0.188)	(0.132)	(0.118)	(0.193)	(0.071)	(0.148)	(0.107)
Social (dummy for employment of former graduate student from the host university)	2.468***	2.373***	1.698**	2.287***	1.587***	2.360***	3.176***	2.939***	1.890***	2.559***
	(0.088)	(0.412)	(0.365)	(0.345)	(0.245)	(0.186)	(0.537)	(0.200)	(0.168)	(0.295)
Attributes of the research group and host university										
Host university incorporated as a commercial enterprise (dummy)	0.772***	1.499	0.473	1.535	1.194	1.288	1.378	0.677*	1.069	1.359
	(0.060)	(0.418)	(0.472)	(0.440)	(0.246)	(0.240)	(0.475)	(0.139)	(0.178)	(0.277)
Age of the research group	1.017***	1.001	1.004	1.014**	1.021***	1.012***	1.008	1.021***	1.020***	1.017***
	(0.001)	(0.008)	(0.009)	(0.007)	(0.008)	(0.002)	(0.009)	(0.002)	(0.002)	(0.004)
Number of researchers of the research group	1.011***	0.988**	0.989	1.003	0.997	1.010***	1.003	1.015***	0.988***	1.011***
	(0.001)	(0.006)	(0.007)	(0.005)	(0.006)	(0.003)	(0.004)	(0.002)	(0.004)	(0.003)
Number of private partners of the research group	1.070***	1.167***	1.134***	1.110***	1.143***	1.092***	1.137***	1.058***	1.137***	1.135***
	(0.002)	(0.012)	(0.018)	(0.012)	(0.013)	(0.005)	(0.013)	(0.003)	(0.006)	(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 presented in UNESCO-UIS (2015). The broad field 'Services' was not estimated due to the small sample size. [b] Location of research groups.
Odds ratios are reported. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.
Source: calculated by the authors, based on CAPES (2017), CNPq (2016), and Ministry of Economics (2015).

Table 4. Estimated parameters of the second stage – decision to collaborate. Logit Model (with robust variance-covariance matrix). Dependent variable: *collaboration*$_i$ (dummy indicating whether the private organization engaged in collaboration).

	Broad Fields of Education and Training[a]									
Independent variables	(1) All fields	(2) Education	(3) Arts and Humanities[b]	(4) Social Sciences, Journalism and Information	(5) Business, Administration and Law	(6) Natural Sciences, Mathematics and Statistics	(7) Information and Communication Technologies	(8) Engineering, Manufacturing and Construction	(9) Agriculture, Forestry, Fisheries, and Veterinary	(10) Health and Welfare
Proximity factors										
Geographical distance (100 km)	0.984*	1.125**	0.487	0.582	0.337***	1.047	0.924	0.912***	1.154***	0.784***
	(0.009)	(0.066)	(0.225)	(0.293)	(0.093)	(0.258)	(0.159)	(0.026)	(0.057)	(0.058)
Institutional (private host university)	0.379***	9.546***	.	3.296	1.855	5.766	0.302	0.107***	0.065***	0.174***
	(0.033)	(4.933)		(6.963)	(4.452)	(10.852)	(0.519)	(0.023)	(0.027)	(0.061)
Social (dummy for employment of former graduate student from the host university)	4.217***	24.302***	.	58.789***	2.568	11.927	30.537***	2.725***	4.392***	5.946***
	(0.321)	(10.363)		(52.244)	(1.541)	(23.370)	(24.093)	(0.559)	(1.078)	(2.538)
Attributes of the research group and host university										
Host university incorporated as a commercial enterprise (dummy)	1.042	2,718.36***	.	3.673	0.000***	679.68**		6.534***	4.170***	1.519
	(0.124)	(5,105.127)		(11.295)	(0.000)	(1,943.205)		(2.185)	(1.811)	(0.769)
Age of the research group	1.014***	0.986	2.491***	0.983	0.952	0.994	1.019	0.986	0.965***	0.977
	(0.002)	(0.028)	(0.744)	(0.103)	(0.167)	(0.049)	(0.060)	(0.009)	(0.009)	(0.015)
Number of researchers of the research group	0.979***	0.962**	.	1.034	1.310***	1.100	1.002	0.992**	0.964**	0.995
	(0.002)	(0.015)		(0.064)	(0.070)	(0.160)	(0.089)	(0.004)	(0.015)	(0.015)
Number of private partners of the research group	1.001	1.407***	.	1.749**	0.764	0.941	1.254	1.001	0.907***	0.984
	(0.004)	(0.097)		(0.407)	(0.132)	(0.105)	(0.295)	(0.008)	(0.031)	(0.046)
Features of the private organization										
Size (number of employees)	1.001***	1.001**	1.117**	1.001**	1.003***	1.001***	1.001***	1.001***	1.001***	1.001***
	(0.000)	(0.000)	(0.049)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Organization incorporated as a commercial firm (dummy)	0.417***	0.312***	.	0.193	0.002***		0.126*	1.301	4.576***	0.474***
	(0.034)	(0.092)		(0.215)	(0.002)		(0.149)	(0.749)	(0.854)	(0.107)
Absorptive capacity: share of employees with an undergraduate degree	8.193***	18.956***	.	0.577	1.533	0.000*	0.805	4.502***	0.572	3.955
	(1.038)	(9.276)		(2.644)	(1.145)	(0.000)	(0.811)	(2.093)	(0.419)	(3.540)
Absorptive capacity: share of employees with a graduate degree	8.368***	5.994***	.	4.232**	73.307***	1.564	8.099***	10.078***	10.799***	4.835***
	(0.412)	(3.434)		(2.973)	(86.407)	(1.064)	(3.213)	(1.136)	(1.407)	(1.008)
Constant	0.000***	0.000***	39.569	0.000***	8,464.397***	0.000*	0.000***	0.000***	0.015***	0.000***
	(0.000)	(0.000)	(233.839)	(0.000)	(26,130.672)	(0.000)	(0.000)	(0.000)	(0.011)	(0.000)
State dummies[b]	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Sector dummies (CNAE 2-digit)	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Chi-squared	12,397	786.7	226.1	920.0	351.6	1,364	286.0	2,490	2,438	922.1
Prob > chi2	0	0	.	0	0	0	0	0	0	0
Pseudo-R^2	0.306	0.338	0.343	0.358	0.679	0.345	0.248	0.287	0.289	0.279
Log-pseudo-likelihood	-16,819	-431.2	-16.38	-65.43	-90.69	-84.67	-301.4	-3,699	-2,136	-898.7
No. of Obs.	2,247,423	120,777	272	21,402	14,837	30,446	37,397	827,943	948,538	148,608

[a] According to the classification presented in UNESCO-UIS (2015). [b] Location of private organizations.
Odds ratios are reported. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: calculated by the authors, based on CAPES (2017), CNPq (2016), and Ministry of Economics (2015).

The estimates for the first stage (Table 3) provide a strong empirical support to *Hypothesis 1*. In all broad fields (and in the entire sample), the estimated social proximity coefficients are positive and statistically significant. This suggests that private organizations are more likely to partner with a research group if it is hosted by a university in which one or more of their employees attended graduate education. Considering the entire sample (column 1), relative odds are 2.5 times higher for the choice of these groups. Although the magnitude of parameters varies per broad field, the likelihood is at least 1.6 higher (odds ratio found for 'Business, Administration and Law').

Following the reasoning summarized in *Hypothesis 1*, we construe this result as an indication of the value of academic relations of former graduate students to private organizations. These social ties help to reduce risks and to facilitate communication and negotiation with universities and research groups, attaching a higher expected value to potential collaborations with them. As a result, private organizations would rather collaborate with socially proximate research groups (all other things being equal), as suggested by the positive value of the social proximity parameters presented in Table 3.

The results for the first stage also confirm that scientific collaboration is spatially concentrated, as concluded by previous studies (Antonelli, 2000; Rybnicek & Königsgruber, 2019). Private organizations are around 21% less likely to choose a research group to partner for each 100 kilometers of distance between their cities (whole sample estimate in column 1). Estimated parameters for different broad fields are all statistically significant and similar, with odds ratio ranging from 0.74 to 0.83. The usual arguments used to explain this result are that knowledge transfer across distance is costly, that local collaboration reduces the risk of loss of information, and that face-to-face interactions facilitate the transmission (and generation) of tacit knowledge, and the development of interpersonal relationships (Autant-Bernard et al., 2007; Fitjar & Gjelsvik, 2018; Hong & Su, 2013; Ponds et al., 2007).

The evidence presented for institutional proximity, on the other hand, is not so strong: for all broad fields a positive association is found, but the parameter is only statistically significant for three of them ('Education', 'Social Sciences, Journalism and Information', and 'Natural Sciences, Mathematics and Statistics'). When considering the entire sample, the estimates indicate that private organizations are 15% more likely to partner with a research group from a private university (a statistically significant result even considering a 0.01 threshold). This topic requires further investigation, as public and private universities in Brazil (and their graduate programs) commonly present very

distinct features on their institutional framework, approach towards scientific activities, focus on different fields of education, and their openness to interactions with industry (Arocena & Sutz, 2005; CGEE, 2016).

Among the attributes of the research group and its host university, we find evidence that age of the group is positively associated with the likelihood of choice for collaboration, although the estimated odds ratios are small (relative odds are 1% to 2% higher for each additional year of existence) and not significant in three broad fields ('Education', 'Arts and Humanities', and 'Information and Communication Technologies'). The number of active researchers is also found to be a significant predictor for the entire sample (likelihood is 1% higher per additional researcher), but the evidence is weak when we break down the estimates by broad field, as only in three of them we find a similar result, and for other two the association is found to be negative and significant.

Table 4 presents the results for the second stage, that models the decision of private organizations to engage in collaboration. Again, we find strong evidence that social proximity is positively associated with this decision. Private organizations are more than four times more likely to engage in collaboration if one or more of its employees have attended graduate education in the university that hosts the most likely partner. This positive association is also found for all but one broad field, and it is statistically significant for six of them.[20] These results provide empirical support to the proposition presented in *Hypothesis 2.* According to the proposed theoretical framework, by reducing the cultural and language gap between organizations, academic relations of former graduate students facilitate the communication and identification of potential projects and technologies to be developed, adding value to the expected return of the collaborative initiative. As the social proximity improves the prospects of a successful and profitable collaboration (with the most likely partner), the private organization becomes more inclined to use this strategy for developing new technologies and find new market opportunities.

The estimated parameter for geographical distance is negative in most cases, as expected, although it is only negative and statistically significant for three broad fields ('Business, Administration and Law', 'Engineering, Manufacturing and Construction' and 'Health and Welfare'). We also found that a greater distance is associated with a

[20] The coefficient for social proximity in 'Arts and Humanities' was not estimated due to the small sample size.

higher likelihood of collaboration in ‘Education’ and ‘Agriculture, Forestry, Fisheries, and Veterinary’. These results require further investigation, as it challenges the widespread evidence in the literature of a negative correlation between distance and collaboration (Rybnicek & Königsgruber, 2019). A possible explanation would be the positive association of distance with innovation value, as geographically distant partners tend to have distinct and non-overlapping knowledge bases, that would favor radical innovation (Petruzzelli, 2011). Also, firms may prefer to collaborate with distant universities if they have higher quality than local institutions (Laursen et al., 2011).

The negative parameter for institutional proximity at the second stage (considering the entire sample) suggests that private organizations are more likely to collaborate if their most likely partner is hosted by a public university. This result needs further investigation, as the estimates of the first and second stages point to different directions.

The estimates for absorptive capacity corroborate previous evidence that organizations with higher capacity are more likely to engage in collaboration (Balland, 2012; Giuliani & Bell, 2005). The results indicate that private organizations are more than 8 times more likely to collaborate per additional share of employees with a graduate degree in its workforce (a similar magnitude is found for share of employees with undergraduate degree). The parameter for this variable is also positive and statistically significant for nearly all broad fields. The main explanation presented in the literature for this association is that a higher ability to absorb and exploit external knowledge increases the benefit expected from the collaboration (Balland, 2012).

But these estimates (for absorptive capacity) are also important for they suggest and present evidence of another channel through which graduate degree personnel may contribute to research collaboration (in addition to their academic relations), and, indirectly, to innovation. This constitutes an additional contribution of this paper, especially considering the “striking absence of systematic research” on the benefits of graduate education in the literature (Halse & Mowbray, 2011). By adding to the absorptive capacity and knowledge base of their organizations, these highly-qualified employees enhance their ability to assimilate knowledge and generate new technology, and henceforth to extract value from research collaborations, making them more interested (and therefore more likely) to collaborate.

Features of both the research groups and private organizations are also found to be important predictors of the decision to collaborate. Larger firms are more likely to engage in collaboration in all broad fields, confirming the association between firm size

and innovation widely acknowledged in the literature (Acs & Audretsch, 1987; Shefer & Frenkel, 2005). Nonprofit organizations also present higher relative odds to collaborate than commercial firms, a result that suggests an interesting research topic, as it signals a higher research orientation of these entities. The age of the research group is found to positively associated with collaboration, following the results of the first stage, while a higher number of researchers is (unexpectedly) associated with a lower likelihood of such outcome.

The main result and contribution of this analysis to the literature is the evidence that academic relations of former graduate students are important predictors of university-firm collaboration. This is an argument that has not been considered nor tested in previous empirical studies that used Boschma's (2005) proximity framework, as reviewed in section 2. As graduate students leave academia to work in private organizations, their personal and professional connections within the scientific community bring organizations closer. We argue that 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. Based on such arguments, we hypothesize and present evidence of a positive association of employing former graduate students with the likelihood of both the choice of partner or formation of a link (*Hypothesis 1*), and of the decision of a private organization to engage in collaboration (*Hypothesis 2)*.

One could argue that our measure of academic relations could also represent a higher cognitive proximity between institutions, as researchers trained in the same university not only have personal ties but are also more likely to have similar knowledge bases within their university (Nooteboom et al., 2007). In this empirical analysis, we choose to control for cognitive distance by limiting the private organization's search for potential partners to a specific knowledge field, while using the social proximity to explain academic relations. But the literature acknowledges that the dimensions of proximity may be overlapping and not possible to be disentangled (Knoben & Oerlemans, 2006). Future research may help to clear this point, by suggesting additional variables and measures to distinguish between the social and cognitive dimensions embedded in academic relations.

An additional contribution is that we find that the magnitude of this association varies substantially per broad field of education, supporting the proposition of Rybnicek

and Königsgruber (2019) that scientific disciplines work as 'moderators' of the social dimension of proximity. Explaining the differences between the parameters estimated for each broad field falls beyond the scope of this paper, and it constitutes the object of a future research agenda. We highlight herein the main points that deserve further investigations: first, the academic relations of graduate students seem to be more important for the choice of partner or formation of a link in 'Information and Communication Technologies' and 'Engineering, Manufacturing and Construction' (as presented in Table 3); on the other hand, in 'Business, Administration, and Law', the social proximity parameter has the smallest estimated value at the first stage, and it does not even achieve statistical significance at the second, a result that hints that, in this broad field, collaborations may be more driven by other features or dimensions of proximity.

The findings presented in this paper suggest new approaches for public policy aimed at promoting university-firm collaboration. In a review of the main policies used for such objective, Cunningham and Gök (2012) identified that they were focused on supporting centers and projects, network initiatives, and schemes to embed academics within organizations (such as industry fellowships). None of the instruments described in their review, however, directly exploited academic relations of former graduate students as a lever to foster partnerships. Our model and results provide a rationale and evidence for alternative (and possibly less costly) policies designed to develop and use such relations as a basis for new or stronger collaborations. Different measures may be envisioned for this purpose, such as incentives for firms to employ Masters and Ph.D. graduates, and initiatives to improve the interaction of universities with their graduate level alumni.

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. There are three main reasons for the generalizability of the main results. First, firms both in Brazil and in other economies are sought for new sources of technological and scientific knowledge to support innovation. In the context of new knowledge-intensive technologies (often associated with the 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 policymakers of any country, and our results suggest important insights that contribute to this goal. Finally, our theoretical arguments do not rely on features of the Brazilian

case, and our empirical results do not seem to be critically dependent on such features. The coefficients estimated for different variables (e.g., geographical distance, absorptive capacity and firm size) are in line with previous studies for other countries (as discussed in section 2), suggesting that the factors associated with the formation of a partnership are generally similar to the ones found elsewhere.

This empirical analysis presents limitations that must be considered for interpretation of results. First, it relies on available data in the original datasets, considering only the cases for which full information is available. Although we expect our dataset to be representative of the collaborations between universities and private organizations in the country, we cannot ensure that non-reported or non-available data is 'missing at random', i.e., not correlated with the explanatory and dependent variables (Hughes et al., 2019). In addition, a number of research choices that limited the sample were necessary to ensure feasibility of the analysis (presented in item 4.2). For these reasons, generalization of the findings requires caution. Future studies that use more complete data and relax some of the assumptions imposed by our research choices may help to overcome these shortcomings and present additional evidence.

Besides these limitations, this study does not aim at evidencing a causal effect between the explanatory and dependent variables. Accordingly, the empirical results only confirm that academic relations are significant predictors of collaboration decisions of private organizations, that we construe as empirical support to the research hypotheses presented. 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.

6 Concluding remarks

Collaborations between universities and the private sector for innovation projects provide important benefits for both partners and for the economy, as a source of technological development, productivity enhancement and industrial competitiveness. Alumni of Master's and Ph.D. programs working at the industry are in a privileged position to foster these partnerships, as they are a point of contact between the scientific and industrial communities, with personal relations in both sides that can help to set goals and objectives, and to reach agreements that are beneficial from both scientific and commercial perspectives.

This paper discusses how academic relations of former graduate students improves social proximity between firms and universities, thus contributing to these collaborations. 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 private organization attended graduate education, the employer organizations is more likely to choose such group to partner (relative odds around 2.5 times higher) and to engage in collaboration (odds ratio more than 4 times higher). Positive and statistically significant associations are found in both stages for the entire sample and for nearly all broad fields. These results are the main contributions of the paper to the understanding of university-firm collaboration, and to the economic geography literature.

The study also presents other important results worth noticing: first, we find that the magnitude of the parameters varies with the broad field of education, in support of the argument that scientific disciplines work as 'moderators' of proximity factors; the study also confirms the spatial concentration of research collaboration, reinforcing the importance of in-person contact for these projects and for transfer of tacit knowledge; it also evidences that absorptive capacity is a predictor of collaboration, pointing to another benefit of graduate degree personnel to industrial innovation and science-industry partnerships; and we also find that larger firms are more likely to collaborate, and that other individual features of both partners are associated with this decision, and they must therefore be considered or controlled for, along with proximity factors.

The analysis also points to future research questions to improve our understanding of the connections between graduate education and university-firm collaboration. The parameters estimated for academic relations in individual broad fields suggest that such ties may play different roles in each one, pointing to the need for specific (and possibly qualitative) studies for each particular field. Also, as previous studies have used other variables to measure social proximity (as discussed in the introductory section), an empirical analysis considering all these variables may provide a clearer picture of how different networks predict university-firm collaboration. Finally, it would be important to replicate this empirical investigation using data from other countries, in order to confirm our statement that the arguments and findings presented herein can be generalized to other economies and industrial contexts.

7 References

Acs, Z. J., & Audretsch, D. B. (1987). Innovation, market structure, and firm size. *The review of Economics and Statistics*, 567-574.

Antonelli, C. (2000). Collective knowledge communication and innovation: the evidence of technological districts. *Regional Studies, 34*(6), 535-547.

Arocena, R., & Sutz, J. (2005). Latin American Universities: From an Original Revolution to an Uncertain Transition. *Higher education, 50*(4), 573-592. doi:10.1007/s10734-004-6367-8

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.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review, 86*(3), 630-640.

Autant-Bernard, C., Billand, P., Frachisse, D., & Massard, N. (2007). Social distance versus spatial distance in R&D cooperation: empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in regional Science, 86*(3), 495-519.

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.

Balconi, M., & Laboranti, A. (2006). University–industry interactions in applied research: The case of microelectronics. *Research Policy, 35*(10), 1616-1630.

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.

Balland, P.-A., Boschma, R., & Frenken, K. (2015). Proximity and innovation: From statics to dynamics. *Regional Studies, 49*(6), 907-920.

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.

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

Boschma, R., & Frenken, K. (2010). The spatial evolution of innovation networks: a proximity perspective. In R. Boschma (Ed.), *The Handbook of Evolutionary Economic Geography* (pp. 120-135). Cheltenham: Edward Elgar.

Broekel, T. (2015). The co-evolution of proximities–a network level study. *Regional Studies, 49*(6), 921-935.

Broekel, T., & Hartog, M. (2013). Explaining the structure of inter-organizational networks using exponential random graph models. *Industry and Innovation, 20*(3), 277-295.

Broström, A. (2010). Working with distant researchers—Distance and content in university–industry interaction. *Research Policy, 39*(10), 1311-1320.

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.

CAPES. (2017). *Discentes da Pós-Graduação Stricto Sensu do Brasil (base de dados confidencial) [Graduate Students in Brazil - confidential dataset]*.

Carayol, N. (2003). Objectives, agreements and matching in science–industry collaborations: reassembling the pieces of the puzzle. *Research Policy, 32*(6), 887-908.

Casey, B. H. (2009). The economic contribution of PhDs. *Journal of Higher Education Policy and Management, 31*(3), 219-227.

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.

Cassiolato, J. E. (2015). Evolution and Dynamics of the Brazilian National System of Innovation. In P. Shome & P. Sharma (Eds.), *Emerging Economies: Food and Energy Security, and Technology and Innovation* (pp. 265-310). New Delhi: Springer India.

CGEE. (2016). *Mestres e doutores 2015: Estudos da demografia da base técnico-científica brasileira [Master's and Ph.D. degree holders 2015: demographic studies on the Brazilian technichal-scientific base]*. Brasília: CGEE.

Cheng, S., & Long, J. S. (2007). Testing for IIA in the multinomial logit model. *Sociological methods & research, 35*(4), 583-600.

CNPq, C. N. d. D. C. e. T. (2016). *Censo do Diretório dos Grupos de Pesquisa 2016*. Retrieved from: 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.

Colombelli, A., Krafft, J., & Quatraro, F. (2014). The emergence of new technology-based sectors in European regions: A proximity-based analysis of nanotechnology. *Research Policy, 43*(10), 1681-1696.

Cosh, A., Fu, X., & Hughes, A. (2005). *Management characteristics, collaboration and innovative efficiency: evidence from UK survey data*: Centre for Business Research, University of Cambridge Cambridge, UK.

Cunningham, P., & Gök, A. (2012). *The Impact and Effectiveness of Policies to Support Collaboration for R&D and Innovation*. Compendium of Evidence on the effectiveness of Innovation Policy. Nesta, Manchester Institute of Innovation Research Manchester. Retrieved from https://media.nesta.org.uk/documents/the_impact_and_effectiveness_of_policies_to_support_collaboration_for_rd_and_innovation.pdf

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.

De Fuentes, C., & Dutrénit, G. (2012). Best channels of academia–industry interaction for long-term benefit. *Research Policy, 41*(9), 1666-1682.

Drejer, I., & Østergaard, C. R. (2014). *The role of geographical, cognitive and social proximity in university-industry collaboration on innovation.* Paper presented at the 9th Regional Innovation Policy Conference.

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.

Fitjar, R. D., & Gjelsvik, M. (2018). Why do firms collaborate with local universities? *Regional Studies, 52*(11), 1525-1536.

Fontana, R., Geuna, A., & Matt, M. (2006). Factors affecting university–industry R&D projects: The importance of searching, screening and signalling. *Research Policy, 35*(2), 309-323.

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.

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.

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.

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.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American journal of Sociology, 91*(3), 481-510.

Greene, W. H. (2011). *Econometric analysis*. Upper Saddle River: Prentice Hall.

Halse, C., & Mowbray, S. (2011). The impact of the doctorate. In: Routledge.

Hanel, P., & St-Pierre, M. (2006). Industry–university collaboration by Canadian manufacturing firms. *The Journal of Technology Transfer, 31*(4), 485-499.

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.

Hossler, D., Braxton, J., & Coopersmith, G. (1989). Understanding student college choice. *Higher education: Handbook of theory and research, 5*, 231-288.

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.

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. doi:https://doi.org/10.1093/ije/dyz032

Isabel Maria, B. F., Rossi, F., & Geuna, A. (2014). Collaboration objectives and the location of the university partner: Evidence from the Piedmont region in I taly. *Papers in regional Science, 93*, S203-S226.

Kingston, P. W., & Clawson, J. G. (1985). Getting on the fast track: Recruitment at an elite business school. *International Journal of Sociology and Social Policy, 5*(4), 1-17.

Knoben, J., & Oerlemans, L. A. (2006). Proximity and inter-organizational collaboration: A literature review. *international Journal of management reviews, 8*(2), 71-89.

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.

Lane, P. J., Koka, B. R., & Pathak, S. (2006). The reification of absorptive capacity: A critical review and rejuvenation of the construct. *Academy of management review, 31*(4), 833-863.

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.

Laursen, K., & Salter, A. (2004). Searching high and low: what types of firms use universities as a source of innovation? *Research Policy, 33*(8), 1201-1215.

Leitch, S. (2006). *Prosperity for all in the global economy-world class skills*: The Stationery Office.

Lin, L., Geng, X., & Whinston, A. B. (2005). A sender-receiver framework for knowledge transfer. *MIS quarterly*, 197-219.

Lindley, J., & Machin, S. (2016). The rising post-college wage premium in America and Britain. *Economica, 83*(330), 281-306. doi:https://doi.org/10.1111/ecca.12184

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. doi: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.

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

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press.

Mertens, A., & Röbken, H. (2013). Does a doctoral degree pay off? An empirical analysis of rates of return of German doctorate holders. *Higher education, 66*(2), 217-231. doi:https://doi.org/10.1007/s10734-012-9600-x

Ministry of Economics. (2015). *RAIS - Annual Social Information Report [Relação Anual de Informações Sociais], confidential microdata*. Retrieved from: http://www.rais.gov.br/sitio/sobre.jsf

Muscio, A. (2013). University-industry linkages: What are the determinants of distance in collaborations? *Papers in regional Science, 92*(4), 715-739.

Nerad, M., & Evans, B. (2014). *Globalization and Its Impacts on the Quality of PhD Education: Forces and Forms in Doctoral Education Worldwide*. Rotterdam: Sense Publishers.

Nooteboom, B. (2000). Learning by interaction: absorptive capacity, cognitive distance and governance. *Journal of management and governance, 4*(1-2), 69-92.

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.

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.

Petruzzelli, A. M. (2011). The impact of technological relatedness, prior ties, and geographical distance on university–industry collaborations: A joint-patent analysis. *Technovation, 31*(7), 309-319.

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

Ponomariov, B. (2009). Student centrality in university–industry interactions. *Industry and Higher Education, 23*(1), 50-62.

Raddon, A., & Sung, J. (2009). The career choices and impact of PhD graduates in the UK: A synthesis review. *Science in Society Programme*.

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.

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.

Sauermann, H., & Stephan, P. E. (2010). *Twins or strangers? Differences and similarities between industrial and academic science* (0898-2937). Retrieved from

Seaman, S. R., & White, I. R. (2013). Review of inverse probability weighting for dealing with missing data. *Statistical methods in medical research, 22*(3), 278-295. doi:https://doi.org/10.1177/0962280210395740

Shefer, D., & Frenkel, A. (2005). R&D, firm size and innovation: an empirical analysis. *Technovation, 25*(1), 25-32.

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

Smith, A., Bradshaw, T., Burnett, K., Docherty, D., Purcell, W., & Worthington, S. (2010). One step beyond: Making the most of postgraduate education. *London: BIS*.

Sorenson, O., Rivkin, J. W., & Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy, 35*(7), 994-1017.

Suzigan, W., Albuquerque, E., Garcia, R., & Rapini, M. (2009). University and industry linkages in Brazil: some preliminary and descriptive results.

Teirlinck, P., & Spithoven, A. (2013). Research collaboration and R&D outsourcing: Different R&D personnel requirements in SMEs. *Technovation, 33*(4-5), 142-153.

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.

Train, K. E. (2003). *Discrete choice methods with simulation*. New York: Cambridge university press.

UNESCO-UIS, U. I. f. S. (2015). *International Standard Classification of Education - Fields of education and training 2013 (ISCED-F 2013) – Detailed field descriptions*. Retrieved from Montreal:

http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-fields-of-education-and-training-2013-detailed-field-descriptions-2015-en.pdf

Vanhaverbeke, W. (2006). The interorganizational context of open innovation. *Open innovation: Researching a new paradigm*, 205-219.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.

Zhang, L. (2005). Advance to graduate education: The effect of college quality and undergraduate majors. *The Review of Higher Education, 28*(3), 313-338. doi:https://doi.org/10.1353/rhe.2005.0030