



HAL
open science

Collaboration networks within a French cluster: Do partners really interact with each other?

Bastien Bernela, Rachel Levy

► **To cite this version:**

Bastien Bernela, Rachel Levy. Collaboration networks within a French cluster: Do partners really interact with each other?. *Papers in Regional Science*, 2017, 96 (1), pp.115 - 138. 10.1111/pirs.12170 . hal-01525808

HAL Id: hal-01525808

<https://univ-tlse2.hal.science/hal-01525808v1>

Submitted on 16 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Collaboration networks within a French cluster: Do partners really interact with each other?

Bastien Bernela¹, Rachel Levy²

¹ CRIEF, Poitiers, France (e-mail: bastien.bernela@univ-poitiers.fr)

² LEREPS, Toulouse, France (e-mail: rachel.levy@iut-tlse3.fr)

Received: 6 May 2014 / Accepted: 24 February 2015

Abstract. We discuss the common hypothesis of complete graph representation according to which, in collaborative projects, all partners interact with each other in homogeneous ways. More precisely, this research aims to determine the heterogeneity in terms of existence and frequency of interactions between dyads of organizations that jointly participated in collaborative projects. From a survey of participants involved in innovation projects approved by a French cluster, we collect information about 754 collaboration ties. Using an ordered probit with selection equation, we then test the impact of several determinants, including geographical proximity, on the existence and frequency of the observed interactions.

JEL classification: L52, O32, R58

Key words: Collaboration tie, interaction, inter-organizational networks, cluster, complete graph

1 Introduction

Innovation is more and more collaborative, both for inter-firm R&D (Hagedoorn and Roijakkers 2006) and for scientific partnerships (Wagner and Leydesdorff 2005). According to Kline and Rosenberg's (1986) model, knowledge creation and innovation can be viewed as collective processes. Focusing particularly on the spatial dimensions of collaboration, recent works have mobilized network analysis as a research tool to better understand how networks function (Boschma and Frenken 2010). A common way to study collaborations for innovation consists in focusing on interactions between partners that form an inter-organizational network. 'More and more researchers get convinced that networks are an appropriate conceptualization of inter-organizational interaction and knowledge flows' (Ter Wal and Boschma 2009, p. 740), information-exchange process being embodied in inter-organizational linkages.

The objective of this paper is twofold. First, we want to discuss the complete graph hypothesis, commonly mobilized in existing empirical studies, according to which every dyad of partners involved in a common innovation project interacts in a homogeneous way. The structure of the overall network can be modified if some ties do not exist, and when existing, the

nature of inter-organizational ties can differ regarding the frequency and the channel (face-to-face vs. information and communications technology (ICT)) of interaction. This paper aims at testing whether complete graph representations of scientific and innovative networks are as accurate as they are assumed to be in the empirical analyses reported in the great majority of the literature. Second, we question the determinants of the existence (or the absence) of interactions between partners that jointly participated to a collaborative funded project, and when existing, the determinants of the frequency of interactions, distinguishing between physical (face-to-face) and non-physical (using ICT) interactions.

To summarize, this research contributes to the existing literature by testing complete graph hypothesis and by investigating the determinants of inter-organizational interactions, using dyads of organizations that jointly participated in a project as units of analysis. As data about existence and frequency of interaction between partners within collaborative projects does not exist, we have conducted an online survey and proposed a case study.

Geography of innovation scholars developed the idea that innovation is a localized process, and particularly takes place within clusters, that is 'geographic concentrations of interconnected companies, specialized suppliers, service providers, firms in related industries, and associated institutions (for example, universities, standards agencies and trade associations) in particular fields that compete but also co-operate' (Porter 1998, p.197). The literature has underlined how this notion of 'the cluster' includes both the concepts of network and of geographical proximity. In their introduction to a special issue of *Papers in Regional Science*, Brenner et al. (2011) underline three challenges in studying knowledge networks within clusters; the third one deals with the methodological challenges involved in better understanding the emergence and the structure of knowledge networks within clusters. In line with this challenge, this paper aims to better understand interaction practices within collaborative projects. Empirically, the data used is taken from a survey among the participants of projects labelled by a French cluster specialized in knowledge-intensive industries.

In France, clusters have mainly been implemented through the government's competitiveness clusters (CC) policy.¹ Empirically, we focus on the case of projects conducted in a specific French CC, collecting data from an online survey addressed to all the partners involved in collaborative projects 'labelled' by this particular cluster. The survey, which was based on a sample of 88 collaborative projects conducted between 2006 and 2012, asked partners involved in innovation projects to assess the frequency of their interactions with every other partner, from which we collected information about 754 ties linking two partners. The survey provides original data, as the nature of the collaboration ties is described by the actors themselves.

The paper is structured as follows. In Section 2, we present the methodological background about innovation collaborative networks building and the determinants of partners' interactions during the length of projects, and introduce our main hypothesis. In Section 3, we present the available data from the cluster, the survey and the collected variables. Then, in Section 4, we give some descriptive statistics about the existence of interaction between partners that jointly participated in innovation project, according to the characteristics of projects and partners. This allows us to discuss complete graph hypothesis, by comparing the underlying theoretical network with empirical one built with declarative interaction from survey data. In Section 5, we test an ordered probit with selection equation to identify the determinants of existence and frequency of interaction between actors involved in a collaborative project. We conclude by discussing methodological and policy issues about networks.

¹ We use the term competitiveness cluster to translate the French term 'pôle de compétitivité', used to designate the French government's cluster support policy (<http://competitivite.gouv.fr/home-903.html>). These clusters aim to reinforce the competitiveness of territories and of cluster members.

2 Empirical issues about network building

2.1 *Identifying collaborative innovation processes*

From the empirical articles about collaborative innovation processes published over the 20 last years, we can identify several types of data used. Interactions between firms and/or research labs within innovation networks can take different forms and can be measured by different indicators, in particular alliances (Gay and Dousset 2005; Stuart et al. 2007; Gilsing et al. 2008), co-authorship in scientific publications (Ponds et al. 2007; Fritsch and Kauffeld-Monz 2010; Hoekman et al. 2010), co-patenting (Hussler and Ronde 2007; Carayol and Roux 2008; Breschi and Lissoni 2009; Hanaki et al. 2010), European Programmes (Breschi and Cusmano 2004; Roediger-Schluga and Barber 2008; Autant-Bernard et al. 2007b) and research consortia (Busom and Fernandez-Ribas 2008; Cassiman et al. 2010; Vicente et al. 2011), or even PhD students co-supervised between science and industry sponsors (Levy 2005; Bouba-Olga et al. 2012). To collect these different data, some authors build their own one, mainly using case studies or surveys (Boardman and Bozeman 2006; Arvanitis et al. 2008; Bekkers and Freitas 2008; Cassiman et al. 2010), while others use existing datasets, including international surveys as the European community innovation survey (Cassiman and Veugelers 2002; Belderbos et al. 2004; Mora-Valentin et al. 2004).

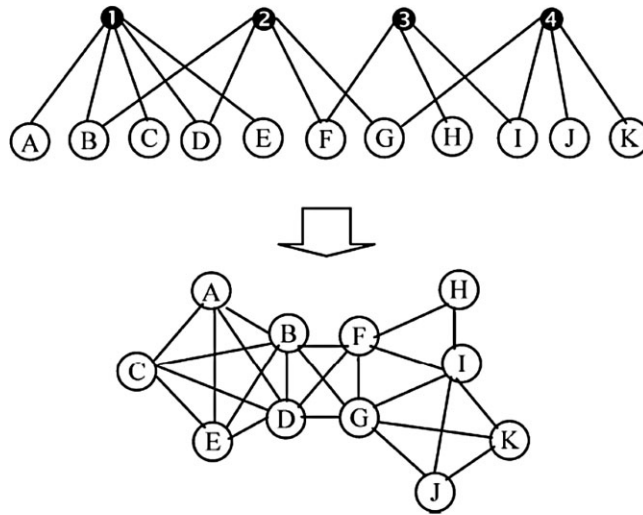
The studies cited above have often specific sectorial and/or territorial delimitations, such as university-industry linkages in Austria (Schartinger et al. 2002), Switzerland (Arvanitis et al. 2008) and France (Ferru 2014), Texas air quality research collaborations (Boardman and Bozeman 2006), New-Zealand biomedical collaborations (He et al. 2009), or European biotechnology (Gay and Dousset 2005; Stuart et al. 2007), nanotechnology (Autant-Bernard et al. 2007b; Cunningham and Werker 2012), IT industry (Hanaki et al. 2010), or GNSS sector studies (Vicente et al. 2011; Balland 2012; Balland et al. 2013). These sectorial and/or territorial delimitations are both the strength and the weakness of case study method, with deep understanding of studied mechanisms but limited possibility to generalize results.

We focus in this paper on the common participation in innovation projects within the framework of a French cluster. While several authors have focused on French CC policy data (Grandclement 2011; Levy and Talbot 2014), there is a significant difference when studying innovation partnerships between focusing on projects that are submitted and those which are actually funded and effectively realized. Using project proposals instead of really implemented projects may be justified when the objective is to identify the determinants of the choice of partners within a project, as Autant-Bernard et al. (2007b). On the contrary, if the objective is to analyse knowledge exchange, as it is the case in this paper, really implemented projects and effective interactions should be considered. Since we aim to measure the existence and frequency of such collaboration, we concentrate in this paper on projects that achieve funding and are actually realized.

2.2 *Moving from raw data to networks: The hypothesis of the complete graph representation*

The studies on innovation collaborative processes noted above have usually been conducted in combination with some recent advance in sociological theory, and particularly social network analysis: indeed, Autant-Bernard et al. (2007a, p. 342) refer to 'the networked nature of knowledge creation and the geography of innovation'. Their objective is to identify, from the network structure and the actors' positions, the best ways to foster innovation (Hussler and Ronde 2007; Balland 2012; Balland et al. 2013).

Independent from the question of the nature of data (as considered in subsection 2.1), we also focus on methodological issues associated with network analysis studies. Indeed, as



Top: Bipartite graph of organisations (A to K) and projects (1 to 4), with lines linking each organisation to the project in which it participated.

Bottom: The one-mode projection of the same network onto just organisations.

Fig. 1. From bipartite to unipartite network

Source: Breschi and Cusmano (2004).

Vonortas (2013, p. 604) recalls, ‘in network analysis the researcher must deal with subtle issues’, the most significant of which concerns unipartite network representations and their construction from empirical data, where he notes that ‘the quality of the results is as good as the data they depend on’. The majority of empirical studies we find in the literature are based on a dominant – and widely accepted – hypothesis that we want to test in our survey: the complete graph representation (Breschi and Cusmano 2004; Autant-Bernard et al. 2007b; Roediger-Schluga and Barber 2008; Balland 2012; Vonortas 2013), which holds that all the partners involved in a collaborative innovation project will interact with each other. Literature often considers ties as supporting knowledge diffusion (Cassiman et al. 2010; Balland 2012), and that the fact of being involved in the same project implies that actors share knowledge (automatic assumptions that Gomes-Casseres et al. 2005 discuss). To build a global network representation from project data, most authors follow Breschi and Cusmano (2004) in transforming bipartite (or 2-mode) network projections – which link actors to the projects in which they are both involved – into unipartite (or 1-mode) projections linking together pairs of actors involved in the same project, as presented in Figure 1. Roediger-Schluga and Barber (2008, p. 326) develop a similar method about R&D projects, and make ‘the assumption that the contract data produces networks that reasonably approximate to actual patterns of interaction’. We could easily expand the list of references to witness the methodological dominance of what has come to be called the complete graph.

Implicit in that representation is the assumption that all actors in a network are connected to all other actors, so that knowledge is automatically diffused throughout the network through their common participation in innovative projects. But, to our knowledge, little is known about real interactions within collaborative projects, so this hypothesis lacks empirical evidence. Consequently, we must pay attention to conclusions driven by social network analyses which apply the complete graph hypothesis to partnership data.

Other works (Breschi and Cusmano 2004; Grandclement 2011) suggest that the co-ordinator of the project (who Breschi and Cusmano call the ‘prime-contractor’) is connected to every

participant through its dominant position, but without observing ties between the other participants, a hypothesis that leads to a star graph representation, and tends to overestimate the strength of ties involving the co-ordinator relative to those involving other partners.

From networks built with this complete graph hypothesis, many scholars propose to characterize network structure and compute indicators of actors' (i.e., nodes') position within a network. More precisely, network structures can be characterized by their size (numbers of nodes and ties), their density (numbers of actual ties divided by the total possible number of ties), their connectivity (number of components, size of the major component, and number of isolated nodes) and geodesic distances (the shortest possible path between two nodes). Finally, there can be indicators about the degree of clustering (presence of 'grapes') within the network, usually measured by the number of transitive triples (Borgatti et al. 2002). For example, Breschi and Cusmano (2004) characterize European program networks by their density, the number of components involved, the size of the two largest components (the core of the network), the degree of clustering, the average and maximum distance between any two nodes, and the average degree centrality of the nodes within the largest component. Roediger-Schluga and Barber (2008) also take into account the entry and exit of actors into and from the network between different periods.

Three indicators of actor centrality are usually employed to measure actors' positions within global networks (Borgatti et al. 2002; Borgatti and Foster 2003): degree centrality (i.e., the number of ties linking a node to other network nodes); closeness centrality (a measure of the distance between one node and other network nodes); and betweenness centrality (a measure of an actor's intermediary position between other network nodes), usually used to measure the level of control exercised by that actor over network activity (Levy and Talbot 2014). Autant-Bernard et al. (2007b) use actor's network positions to measure the social distance between them.

As recalled previously, all these indicators (including network structures and node positions) are calculated under the complete graph hypothesis. To improve the method of network representation, some authors have introduced measures to value the intensity of the ties between different network actors. For example, Roediger-Schluga and Barber (2008) define tie intensity (or weight) as the number of projects in which two partners are involved together: while this methodology can be used to value ties between actors, it does not measure the frequency of their interactions within a specific project, but only how often the partnership has been renewed over time. To our knowledge, Paier and Scherngell (2011) are the only ones to propose a declarative measure of collaboration intensity through a survey toward European program partners.

In this paper, we propose to approximate intensity of interactions by measuring the frequency of interactions between network actors using survey data (we discuss the use of this proxy in subsection 3.3). Thus, the hypothesis we test in this paper is that: interactions between partners involved in a collaborative project are heterogeneous (in terms of existence and frequency). We argue that, in reality, interactions between such partners can be better represented through an empirical graph – since, in practice, some ties exist and some do not, and some ties are stronger than others. Figure 2 opposes the two types of representations: the theoretical (the complete graph) on the left, and the empirical on the right.

2.3 Interacting to exchange knowledge, what determinants?

In the knowledge-based economy, innovation is driven by the interaction of producers and users in the exchange of both codified and tacit knowledge (OECD 1996). According to Reinold et al. (2013, pp. 4–5), 'generation of knowledge is a social process and, therefore, the performance of an economy to generate knowledge crucially depends on successful co-operation between involved actors not only on the individual, but also on the organizational level'. This

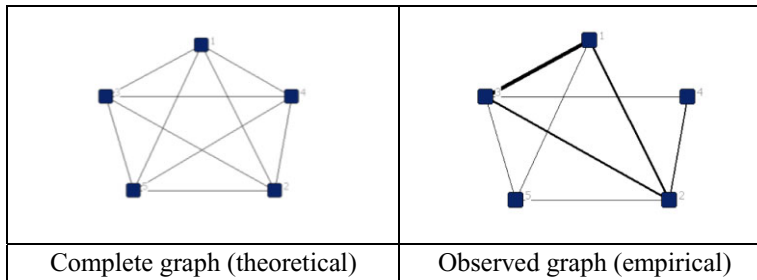


Fig. 2. Representation of theoretical and empirical graphs

co-operation goes through knowledge exchange, and ‘a great part of the necessary knowledge is combined by co-ordinating people to build up inter-organizational capabilities for knowledge generation’ (Reinold et al. 2013, pp. 4–5).

Let us consider the determinants of partner interaction as identified in the literature. Numerous authors have tried to identify the factors which might favour collaborations between innovation actors, whether science-industry linkages or inter-firm collaborations. Many of the studies previously cited – working on patents, publications or common participation to European Framework Programs (Autant-Bernard et al. 2007b; Hussler and Ronde 2007; Ponds et al. 2007; Carayol and Roux 2008; Gilsing et al. 2008; Levy et al. 2009; Fritsch and Kauffeld-Monz 2010; Hoekman et al. 2010) – use social network analysis to test the impact of different forms of proximity on innovation activities. Indeed, it is widely assumed that proximity between partners – whatever its form or definition – has a positive impact on their likelihood to interact and to innovate. As mentioned by Paier and Scherngell (2011, p. 91), ‘a fundamental question addressed by the research on knowledge flows – and hereby we assume that they are sufficiently well captured in inter-organizational R&D collaborations – is how these flows are influenced by spatial, institutional and other forms of proximity’.

Concerning different forms of proximity, despite the wide diversity of proximity grids that have been developed (Boschma 2005; Torre and Rallet 2005; Bouba-Olga and Grossetti 2008; Boschma and Frenken 2010), researchers generally agree that a basic distinction can be made between spatial and non-spatial proximity. Authors from the French school of proximity distinguish precisely between geographical and organizational proximity (Kirat and Lung 1999; Gilly and Torre 2000; Torre and Rallet 2005; Bouba-Olga and Grossetti 2008).

Geographical proximity refers to the spatial separation and the links in terms of distance, and can be measured by physical distance or by localization in the same administrative territory (see Cunningham and Werker 2012). Several studies (Katz 1994; Boschma 2005; Maggioni and Uberti 2009; Scherngell and Barber 2009; Paier and Scherngell 2011) provide evidence that geographical distance between two actors significantly reduces (i) the likelihood they engage R&D collaboration and (ii) the interactions between partners, namely face-to-face communication.

By testing the role of geographical proximity on the existence and frequency of interactions, we will give some evidence on whether being co-located is necessary for a better co-ordination in innovation collaborative projects. Nevertheless, we can expect a different effect of geographical proximity on physical and non-physical interactions. Indeed, two main channels of knowledge diffusion and co-ordination are traditionally distinguished in innovation literature: face-to-face interactions characterized by physical meetings vs. interactions using ICT as phone, videoconference, emails, etc. Note that face-to-face interactions are accompanied by geographical proximity (at least temporary), while non-physical interactions can be done either in geographical proximity or at a distance. According to Rodriguez-Pose and Crescenzi (2008,

p. 55), ‘face-to-face contacts do not only act as a communication technology but also pursue other functions (such as greater trust and incentives in relationship, screening and socializing, rush and motivation) which make communication not only possible but also more effective, and ultimately ease the innovation process’. However, ICT cut the cost associated with knowledge and information transfer through space and reduce spatial constraints, especially for innovative activity (Cairncross 1997; Aguilera and Lethiais 2011).

In terms of non-spatial proximities, the French school defines organizational proximity as ‘the economic separation and links in terms of production organizations’ (Gilly and Torre 2000, pp. 12–13), with the idea that having the same organizational structure facilitates partners collaboration. The nature of partnership can be measured by separating small and medium-size enterprises (SMEs) from big firms (Levy and Talbot 2014), or by distinguishing between science-industry linkages and intra-industrial links (Cunningham and Werker 2012), suggesting collaboration between the scientific sector and the private sector is often difficult since the two sectors pursue different goals. Ponds et al. (2007) demonstrate that institutional proximity between actors can act as a substitute of geographical proximity. The need to be co-located when partners share the same language and culture is reduced, and non-physical interactions through ICT might replace face-to-face ones.

Finally, the existent literature underlines the role of acquaintanceship and social embeddedness (Granovetter 1985) as drivers of interactions: in this context, Boschma (2005) uses the concept of social proximity to refer to the climate of trust between actors that can facilitate collaboration. Other studies (e.g., Gulati 1995; Hagedoorn 2006; Thune 2007; Ferru 2014) have demonstrated empirically the importance for innovation projects of the reactivation of previous collaborations. According to Paier and Scherngell (2011), repeated interactions and long-term relationships encourage mutual learning and a high degree of trust between partners.

In the empirical part of this paper, we provide an econometrical analysis of the determinants of existence and frequency of interactions between partners. The aim is to go beyond debating the complete graph hypothesis, and better understand why interactions within collaborative projects are not homogeneous. We propose to test these three types of determinants – geographical proximity, nature of partnership and acquaintanceship – on the existence and the frequency of interactions between collaborators (see subsection 3.4).

3 Data and method

3.1 Case study on a French competitiveness cluster

In 2005, the French government implemented a national bottom-up cluster policy to create Competitiveness Clusters (CCs), which it defined as ‘joint theme-based initiatives for a given geographical area, i.e., in a given territory, that bring together companies, research centers and educational institutions in order to develop synergies and co-operative efforts targeted at one (or more) given market(s) (. . .) clusters using synergies and innovative joint projects to give their members a chance to be national and international leaders in their fields’ (www.competitivite.gouv.fr). Referring to the objectives of EU Framework Programmes which are ‘to strengthen the scientific and technological bases of European industry and to enhance its international competitiveness’ (Reinold et al. 2013, p. 204), we notice that French CC policy has the same ambition at an inferior geographic scale. Scientific interests about innovation processes are thus the same.

In concrete terms, 71 CCs have been established within French territories (some of them globally oriented, some nationally oriented), each specialized in a sector such as electronics, biotechnology, wood industry, etc. Cluster members are usually located within the same NUTS 2 region, but occasionally spread over two or three contiguous regions. These clusters are all

organized as associations, with memberships that include several firms and research laboratories or higher education and research establishments (HEREs) located in their geographical areas and more or less concerned with their specialized sector or technology.

As well as their management and territorial marketing activities, CCs are also required to encourage the development of innovation projects, especially between cluster members, although they also often involve partners from beyond the clusters' home zones. Firms and HEREs from each cluster propose innovation collaborative projects that are launched in a two-step procedure. First, they are labelled by the CC, depending on the project's innovative characteristics and on its links with the cluster's strategy. In some cases – depending on the project's subject – they may be co-labelled by several CCs, following the second phase of the government's cluster policy (from 2010), which emphasized inter-clustering and co-operation between members of different CCs.

Once labelled, each project must find funding, which can come from various different sources: two national funding schemes: (i) the first administered by the National Research Agency; and (ii) the second from a governmental fund dedicated to CC projects; or (iii) European funding, generally through European Framework Programmes and European Regional Development Fund; or (iv) local funding, mostly from local authorities (regional councils, public investment banks, etc.).

Our case study uses data about projects labelled by a national CC which have been run since the CC policy was implemented in 2005. This CC is specialized in a high technology sector – microwaves, photonics, secure networks, images and digital interfaces – and localized in a small French region.² By the end of 2012, this cluster had acquired 76 members and had labelled 284 projects: comparing these figures with those of others French CCs, this cluster is about average, and so represents a relevant setting in which to analyse collaboration within clusters (see EuroLIO 2010 for a comparison of a sample of 20 similar CCs in terms of numbers of establishments involved and of their employees). Even if the case study methodology needs to be completed by sectorial or regional comparison, it has been largely used notably in the field of cluster studies.³

Among these 284 labelled projects, we focus only on projects which actually gained funding, and which have two or more partners. More precisely, we exclude the 174 projects that failed to attract funding, as some of them may never have been conducted. We also exclude 22 projects that only had a single participant (usually start-up creations), because they do not fit with our aim to study collaboration processes.

This research is therefore based on a sample of 88 collaborative projects conducted between 2006 and 2012 and involving 262 different establishments (firms or HEREs), each of which participated (on average) in 1.8 project (participations per establishment ranged from 1 to 22). In other words, the data we collected represents a total sample of 475 project participations.

3.2 Characteristics of projects and partners

Table 1 describes the data used in this empirical research, detailing some descriptive statistics about participation in projects, some of which relate to the projects, and others to the partners involved in them.⁴ We use the following information to characterize projects:

² This CC wants to remain anonymous, so this article does not refer to anything that could allow it to be identified.

³ For example, some case studies have been detailed in the special issue of *Papers in Regional Science* on regional innovation systems, clusters and knowledge networking, published in 2011, cf. Brenner et al. (2011); others specific territorial and sectorial studies are listed in subsection 2.1.

⁴ Table 1 includes a representativeness test of the survey's respondent population that we comment on in the following section.

Table 1. Descriptive statistics on population and respondents

| | | Targeted population | | Respondent population | | Chi2 test |
|--------------------------------|-----------------|---------------------|-------|-----------------------|-------|-----------|
| | | n | % | n | % | |
| Partner characteristics | local | 166 | 34.9 | 91 | 48.9 | *** |
| | member | 161 | 33.9 | 95 | 51.1 | *** |
| | coordinator | 88 | 18.5 | 51 | 27.4 | *** |
| | structure | | | | | n.s. |
| | HERE | 251 | 52.9 | 112 | 60.2 | |
| | group | 77 | 16.2 | 27 | 14.5 | |
| | SME | 147 | 30.9 | 51 | 25.3 | |
| Project characteristics | project_size | | | | | ** |
| | ≤5 | 169 | 35.6 | 84 | 45.2 | |
| | >5 | 306 | 64.4 | 102 | 54.8 | |
| | funding | | | | | n.s. |
| | cc_policy | 224 | 47.2 | 97 | 52.2 | |
| | research_agency | 135 | 28.4 | 50 | 26.9 | |
| | europe | 46 | 9.7 | 9 | 4.8 | |
| | local | 70 | 14.7 | 30 | 16.1 | |
| | period_label | | | | | n.s. |
| | period1 | 252 | 53.1 | 100 | 53.8 | |
| | period2 | 223 | 46.9 | 86 | 46.2 | |
| colabelling | 173 | 36.4 | 72 | 38.7 | n.s. | |
| Total number of participations | | 475 | 100.0 | 186 | 100.0 | |

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; n.s.: non-significant.

- project size (*project_size*). It is defined by the number of partners involved. The projects studied had between 2 and 19 partners, with a mean of 5.4 partners. In what follows, we use this mean to distinguish two sizes of projects to simplify our analysis: small projects (with a maximum of 5 partners) vs. large projects (with 6 partners or more);
- funding (*funding*). As noted above, we can distinguish four forms of funding: two national forms: that is, from the National Research Agency (*research_agency*) and from the governmental fund dedicated to CC policy (*cc_policy*); as well as European funding (*europe*); and local funding (*local*);
- co-labelling (*colabelling*). We record information about the co-labelling of a project, namely, when it is approved by at least two CCs; and
- year of labelling (*period_label*). The 88 collaborative projects we study were labelled between 2006 and 2012. As with project size, we simplify our analysis by distinguishing two distinct periods: the first phase of the CC policy (*period1*) refers to projects labelled between 2006 and 2009, and the second (*period2*) to those labelled between 2010 and 2012.

We also use some data characterizing the partners participating in the sample projects:

- co-ordinator (*co-ordinator*). Each project is led by a co-ordinator, the establishment that is the driving force behind the project and ensures the smooth running of the collaboration;
- local. We consider that a partner is local when located in the CC’s own area;
- member. Establishments can participate in CCs without actually being cluster members, so we record information about the cluster membership of each project participant; and
- structure. We distinguish three types of structures: HEREs, SMEs (<250 employees), and groups (or larger establishments).

3.3 *Measuring partners' interactions within a collaborative project: a survey method*

The objective of this paper is to provide a critical assessment of the dominant theoretical hypothesis about the nature of collaboration ties within networks. In order to get qualitative and declarative data about participants' actual interactions during real-life projects, we thus decided to conduct an online survey (Eisenhardt 1989; Yin 2003). As Ter Wal and Boschma (2009) report, though databases of existing links (as EU-FP, strategic alliances, patents, etc.) often do not contain further details on the links or the partners involved, the survey-based nature of methodology provides opportunities to gather additional information on the population that might otherwise be unknown. The survey was addressed to all partners involved in the 88 targeted collaborative projects labelled by the CC (i.e., 475 participations), and sent to their referents as noted in the CC's mailing list. We asked them to answer from the framework of a specific project and describe their interactions with all other project partners. The survey was sent by email in early June 2013, and two follow-up emails were sent to non-respondents after a two-week interval. The CC director sent further follow-up emails to cluster members who had still not replied, and we finally closed the survey in mid-July.

Before studying the response rate of the survey and the representativeness of the respondent sample, we give some information about the content of the survey and some methodological justifications. To try to achieve a high response rate, the survey was very short: in fact, it contained only two questions. Bearing our hypothesis in mind, the first question concerned the frequency of the respondent's interactions with all other project partners.

Following a similar method, Reinold et al. (2013) conduct survey to estimate intensity of interactions between partners jointly participated to an innovation project. These authors made the choice of a dichotomous variable, which refers to the subjective impression on the importance of partnership within dyadic relations. We use the same method concerning the unit of analysis, which are pairs of organizations involved in a common project, but we privilege a more objective and quantitative measure of frequency of interaction as a proxy for the strength (and so the intensity) of the interaction and the knowledge flow due to that interaction. According to Ahuja (2000, p. 430), 'sustained interaction is characterized by frequent communication'. Therefore, a method to characterize interaction consists in assessing the frequency of partners' interaction;⁵ null frequency suggesting the non-existence of interaction at all. Ter Wal and Boschma (2009, p. 747), who analyse the stakes of social network analysis as a tool for economic geography, report that survey methodology offers the opportunity to ask for several characteristics for each of the links. Among them, the frequency of interaction may 'serve as an input for a valued graph, in which each link is provided with a strength'.

Following quantitative survey methodology (Eisenhardt 1989; Yin 2003), we proposed a scale of five different frequencies of interaction, as well as of the absence of interaction:

- 0: No interaction at all;
- 1: Very few interactions, i.e., less than once a year;
- 2: Few interactions, i.e., more than once a year but less than once trimester;
- 3: Regular interactions, i.e., more than once a trimester, but less than once a month; and
- 4: Very regular interaction, i.e., more than once a month.

The choice of this scale was motivated by the fact that all projects lasted at least one year, so each possible answer would make sense over that timescale. To ensure the scale was reliable, we

⁵ This hypothesized positive correlation between frequency and intensity of interactions can be contradicted by weak ties literature (Granovetter 1983): few interactions can be only administrative discussion with the co-ordinator of the project without any knowledge diffusion but it also can be more intense exchange of strategic knowledge. Considering our data, we cannot control empirically the existence of weak ties.

discussed and validated it with the CC director. This question was subdivided as follows: partners had to distinguish, with the same frequency scale, physical interactions (face-to-face meetings) and non-physical interactions (using ICT as phone, videoconference, emails, etc.). As recalled previously, this distinction is crucial namely for commenting the geographical dimension of innovation processes.

We then added a second question about the partners' acquaintance before the project, asking if they knew each other and if they had worked together before, with the object of collecting relevant empirical material to build a variable about acquaintanceship (see subsection 3.4).

Of the 475 project participants, we actually sent 371 surveys, as 104 referent email addresses were missing, and finally collected 186 responses, that is a 50 per cent response rate of surveyed partners (and 39% of those initially targeted), which is satisfactory. We ran chi-2 tests on each descriptive variable to check bias relative to all responses to find out whether the respondents' profile differed from those of the overall targeted population (see Table 1). We found that local partners, CC members and project co-ordinators were comparatively overrepresented in the sample of respondents.⁶ We can legitimately assume that these actors were more receptive to our research because of their stronger link to the CC: moreover, the CC director sent follow-up emails to cluster members, which probably increased their response rate. The number of partners involved in projects also appears to have impacted on their likelihood to answer the survey, as the response rate for small projects was higher. Apart from this overrepresentation – which does not modify the interpretation of our results – there was no bias in responses according to respondents' type of structure, or funding, or period of labelling, or co-labelling. We also note that Table 1 shows that HEREs are overrepresented in the population studied in comparison with other clusters (EuroLIO 2010) – and as a consequence – the proportion of projects funded by the National Research Agency is also greater.

3.4 *The particular unit of analysis: The tie between two partners*

To analyse the interactions between partners in innovation projects, we focused at the level of the tie between two partners participating in the same project, and tested the impact of the characteristics of projects and of partners on the existence and frequency of their interactions. This choice of tie analysis requires prior disambiguation – since, for the same tie, we can get two answers describing the level of interactions. For instance, if actors A and B are involved in a same project, A can describe its perception as to its level of interaction with B, and B of its interaction with A. Thus, there are two possible responses about the same tie, and these two responses may differ. In terms of the existence of interactions, of the overall total of 197 ties for which we received two answers (from A and B), in 3.5 per cent of the cases one party reported zero interaction, while the other reported interaction at some level. Concerning the frequency of interactions, we got different answers from the two partners for about 62.4 per cent of the ties, but in most cases (78.1%) those differences represented offsets of only one degree of frequency. Heterogeneous interpretations about interaction frequency from the two partners in same tie are thus low. When the responses from the two partners differed, we used the lower frequency as our measure for that variable. We justify this choice in the following way: as our research question concerns the issue of collaboration intensity, we consider that, when two partners have different perceptions of the intensity of their collaboration, taking the greater perceived intensity into account would risk overestimating the real intensity.

⁶ These three variables were correlated (at 0.1%): more than 90 per cent of members were local, and coordinators were members in 45 of the 88 projects surveyed. We have therefore chosen to focus on the *coordinator* variable and excluded *local* and *member* from our analysis.

Table 2. Descriptive statistics on ties

| | | n | % |
|-------------------------|-----------------|-------|------|
| Project characteristics | project_size | | |
| | ≤5 | 172 | 22.8 |
| | >5 | 582 | 77.2 |
| | funding | | |
| | cc_policy | 250 | 33.2 |
| | research_agency | 350 | 46.4 |
| | europe | 98 | 13.0 |
| | local | 56 | 7.4 |
| | period_label | | |
| | period1 | 366 | 48.5 |
| period2 | 388 | 51.5 | |
| colabeling | 351 | 46.5 | |
| Ties variables | coordinator | 270 | 35.8 |
| | geo_proxi | 217 | 28.8 |
| | partnership | | |
| | sc_sc | 207 | 27.5 |
| | ind_ind | 198 | 26.3 |
| | sc_ind | 349 | 46.2 |
| | acquaintance | 539 | 71.5 |
| Total | 754 | 100.0 | |

The 186 responses to the survey represent 754 different ties. Table 2 describes the composition of this analysis sample using the characteristics of projects defined in subsection 3.2 (size and funding of the project, period of labelling and co-labelling), as well as the presence of the co-ordinator of the project as a control variable. Using the tie as the unit of analysis also allows for the introduction of the following explanatory variables to better define the relationship between the two partners.

3.4.1 Geographical proximity (*geo_proxi*)

Following the methodology proposed by Ferru (2014) and Levy and Talbot (2014) in other French case studies, we build a binary variable to define geographical proximity: two partners are considered geographically close if they are located in the same or the neighbouring (NUTS 3) area. In terms of French geography, this criterion is relevant at the scale of the NUTS 3, the NUTS 2 region being significantly larger. We use the place where the effective project work took place, rather than its administrative headquarters, as the partner's geographical location.

3.4.2 Nature of the partnership (*partnership*)

Following Levy et al. (2009), we distinguish between collaborations where the partners are public and private actors. From data about the type of structure provided by the CC (HEREs, SMEs, groups), we differentiate three types of partnerships. First science-science ties (*sc_sc*) involve two HEREs; second, industry-industry ties (*ind_ind*) involve two SME(s) and/or group(s); and third, science-industry ties (*sc_ind*) involve a HERE and an SME or group.

3.4.3 Acquaintanceship (*acquaintance*)

Previous partnership is a crucial point when studying collaboration, but this information is often the most difficult to assess. One of the main interests of our research is to survey partners

Table 3. Descriptive statistics and binary probit results on the existence of interactions

| | | Descriptive statistics | | Binary probit (n = 754; pseudo $r^2 = 0.20$) |
|-------------------------|-----------------|----------------------------------|--------------------------------|--|
| | | Number of observations (ties) | Percentage with interaction | Marginal effects (dF/dX) |
| Project characteristics | project_size | | | |
| | ≤5 | 172 | 95.3 | ref. |
| | >5 | 582 | 93.1 | -0.1 |
| | funding | | | |
| | cc_policy | 250 | 95.6 | ref. |
| | research_agency | 350 | 97.1 | 0.4 |
| | europe | 98 | 78.6 | -14.9*** |
| | local | 56 | 89.3 | -9.5* |
| | period_label | | | |
| | period1 | 366 | 94.8 | ref. |
| period2 | 388 | 92.5 | 3.3* | |
| colabeling | 351 | 93.2 | -0.9 | |
| Ties variables | coordinator | 270 | 98.5 | 5.3*** |
| | geo_proxi | 217 | 93.6 | -1.8 |
| | partnership | | | |
| | sc_sc | 207 | 96.6 | ref. |
| | ind_ind | 198 | 88.9 | -3.2 |
| | sc_ind | 349 | 94.6 | 0.2 |
| | acquaintance | 539 | 96.3 | 4.1** |
| | Total | 754 | 93.6 | |

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

directly about their previous acquaintance and to test the impact of this variable on interactions within collaborative projects. We asked each partner to report if they knew or had previously worked with the other before the focal project, and used their answers to measure acquaintance-ship as a binary variable, as Paier and Scherngell (2011) did.

4 Results about the existence of interactions between partners

4.1 Descriptive statistics about the existence of interactions and their determinants

One of the main objects of this paper is to empirically test the dominant representation of collaboration networks, called the complete graph – and its underlying general hypothesis that each partner in a collaborative project interacts with every other partner – by simply looking at the existence of interactions between all the partners of innovative projects in our case cluster. We define the existence of interaction as a binary variable as follows:

- interaction between two partners exists if and only if physical interaction and/or non-physical interaction exists (one non-null frequency); and
- interaction between two partners does not exist if and only if both physical interactions and non-physical interactions do not exist (null frequencies).

In the present section about the existence of interactions, results will not involve distinction between physical interactions and non-physical interactions; Section 5 about frequency will deal with this. Table 3 shows that some ties (48 out of 754) are characterized by the absence of any

interaction between partners, but that the great majority (93.6%) of ties involve interactions, a result that tends – at least partly – to support the complete graph hypothesis. However, as the following section shows, these non-existing ties can change some interpretations about innovation network structures. Table 3 gives some descriptive statistics about interactions and the results of a binary probit⁷ test showing the impact of explanatory variables on the existence (or not) of interactions between two partners.

Table 3 identifies four main results. First, we observe the negative and significant impact of European funding on the existence of interactions: more than one fifth of such projects (21.4%) are characterized by the absence of interactions between partners. European projects are often run as sets of sub-projects linked together through a co-ordinator (Breschi and Cusmano 2004), implying that partners do not interact directly with every other. To our knowledge, this result has not been previously noted in the literature, and calls for some restraint *vis-à-vis* the use of the complete graph hypothesis in analysing large European projects.

Second, Table 3 highlights the significant role of the co-ordinator: respondents reported no interactions between the two partners in only 1.5 per cent of ties involving co-ordinators. The co-ordinator variable is highly significant in explaining the existence of interactions between partners, and this marginal effect indicates that being project co-ordinator increases the probability of interacting with other participants by 5.3 per cent. This result demonstrates the central position of co-ordinators, and confirms partially the legitimacy of the star graph representation; while the likelihood that interactions take place between co-ordinators and other partners is very high, there are also many ties between the other partners. As a consequence, we can consider the empirical graph as an intermediary between the complete and the star graph representations: interactions exist in the ties between most partners involved in a collaborative project, and particularly in those involving the co-ordinator.

Third, an interesting result concerns the positive and significant impact of previous acquaintance between actors on the probability they interact: partners who have experience of collaborating together (whatever the form of their previous collaboration) are more likely to interact during their focal project (that confirms results from a similar study of Paier and Scherngell 2011). Of the 754 ties recorded, 539 (71.5%) were characterized by previous acquaintance between partners, which underlines the importance of social proximity as a modality of linkage and as a determinant of the likelihood of future collaborations between the partners (Gulati 1995; Grossetti and Bès 2003; Boschma 2005; Hagedoorn 2006; Thune 2007; Ferru 2014).

Fourth, nor did geographical proximity of partners impact on the probability that they would interact, echoing results by Bouba-Olga et al. (2012) who found a mitigated impact of geographical proximity on collaboration (depending on regional characteristics), and Cunningham and Werker (2012) who also found that geographical proximity (as measured by partners' presence in the same administrative region) had a non-significant impact on the likelihood to interact. On the contrary, Paier et al. (2011) find a significant impact of geographical distance on the existence of collaboration; nevertheless they used a larger definition of geographical proximity than in our study that may explain these two different results. Another possible explanation is the existence of temporary geographical proximity between the two partners (Torre 2008, 2011), allowing partners who were not actually co-located to meet once or twice during collaborative projects. We will test more in depth the effect of geographical proximity by distinguishing physical and non-physical interactions in the frequency analysis.

Finally, several of the variables we have included in our analysis seem to have no significant impact on the likelihood of observing interactions between partners: this is particularly the case

⁷ We test a simple binary probit: the variable explained is the existence of an interaction, which takes the value of 1 if interaction exists, 0 otherwise.

Table 4. Structural characteristics of theoretical and empirical graphs

| | Complete graph (theoretical) | High hypothesis graph (empirical) | Low hypothesis graph (empirical) |
|--|---------------------------------|--------------------------------------|-------------------------------------|
| Number of nodes | 262 | 262 | 241 |
| Number of isolated nodes | 0 | 0 | 21 |
| Size of the biggest component | 100% | 100% | 91.29% |
| Number of ties | 1,335 | 1,282 | 675 |
| Density | 1.95% | 1.87% | 1.97% |
| Average geodesic distance of the biggest component | 3.022 | 3.107 | 3.3 |
| Maximum geodesic distance of the biggest component | 6 | 7 | 8 |

for project size⁸ and co-labelling, but they can impact on the frequency of interaction as we will see in the fifth section. Another variable that appears to have no impact on the probability of partner interaction is the nature of the partnership,⁹ namely, the type of partners involved. Being from different worlds (from science or industry) seems to have no significant impact on how smoothly the project runs, a result that also confirms the findings of Cunningham and Werker (2012) and Reinold et al. (2013).

4.2 Comparison of theoretical and empirical networks

In the previous section, we identified the number of ties without interaction, which is quite low (6.4%). We now aim to assess the impact of these few missing ties on the network properties, comparing structural characteristics and positions of actors between theoretical and empirical graphs.

As explained in subsection 3.3, we could not collect information about the (non-) existence of interactions between partners across the whole network, since our final response rate was 39 per cent of participations. Not having a 100 per cent response rate means we cannot build a full empirical graph representation for the whole network, and we therefore introduce two hypotheses in order to build two estimates (a lower and a higher) of partner interaction rates from the 754 ties for which we did get answers (which represent 56.5% of the 1,335 ties of the complete graph). In our low hypothesis, we suppose that ties between non-respondent partners are characterized by an absence of interaction: only ties for which we get a positive answer confirming the existence of interactions are represented. In contrast, in our high hypothesis, we suppose that interactions did exist between non-respondent partners: so only ties where the actors involved confirmed the non-existence of interactions in their answers are regarded as non-interactive. We propose comparing the complete graph with these two empirical graphs.¹⁰ Table 4 presents some indicators that compare the structure of these three graphs.

In Table 4, we observe no significant differences in terms of density and average geodesic distance between the three graphs, which confirms the results presented in the previous section, and tends to validate the complete graph hypothesis for representing the network of actors inside a cluster. Nevertheless, we note a difference in the graphs' diameters (maximum geodesic distance), which underlines how the presence of some weak ties increases the connectivity of the network (Friedkin 1982; Granovetter 1983).

⁸ We also tested the size as a continuous variable, and it is not significant either.

⁹ We also tested this variable by distinguishing SME from group, and again it is not significant.

¹⁰ Note that survey gives information on partners that have not answered the survey: if actor A and B are involved within the same project, and even if A has not answered the survey, we get information on its interaction with B through B's answer. On the 262 nodes represented on complete graph, 89 are respondents, 152 are non-respondents cited, and only 21 are non-respondents non-cited (that correspond to the 21 isolated nodes in low hypothesis graph). In other words, we get information (directly and indirectly) for 92% of the 262 nodes.

As well as looking at the network structures, we can also compare the three graphs by looking at the position of actors inside the network. Table 5 shows the top ten network nodes in terms of normalized centrality following the three classical modes of calculating centrality as previously defined (degree, closeness, and betweenness) between the two.

We use a Kendall rank correlation test to check whether the differences observed between the complete graph and the two empirical graphs are significant: results reveal they are not, whichever mode of calculation is used, confirming the partial validity of the complete graph hypothesis. Nevertheless, although the differences are not significant, we do observe some differences between the three rankings, particularly concerning betweenness centrality, the position that allows actors to control knowledge diffusion across the innovation network (see for example Levy and Talbot 2014). While the two most central actors are the same for the theoretical and the high hypothesis graph, actor A49 (a technology transfer center located in the same administrative region as the CC) is in fifth position for the complete graph, but is the third most central actor in the high hypothesis graph and first in the low hypothesis graph. Thus, including non-existing ties when applying the complete graph hypothesis decreases the intermediary role of this actor, who is responsible for transferring technology between partners.

These results about the different rankings in nodes' positions, as well of the difference in the network diameters, confirm the idea that even if the complete graph hypothesis is partially acceptable we must be cautious, as the deletion of some ties can change the characterization of the network, and thus the knowledge diffusion within it. This result reinforces the necessity of comparing complete graph with other types of representation in order to discuss the empirical choice of graph representations before drawing hasty conclusions.

5 Results about the frequency of interactions between partners

5.1 *The model*

We have shown that interactions within collaborative innovation projects are not perfectly homogeneous in terms of existence. Table 6 reports partners' answers about their interactions and also demonstrates that interactions are not homogeneous in terms of frequency either. In Table 6, as well as in the rest of this section, we distinguish physical and non-physical interactions; for a same tie, one of the two types of interaction can exist without the other (contrary to the definition in subsection 4.1. where we focus only on the existence of interactions). If the differences are not so important considering the existence of interaction between physical interactions (8.5% of ties without face-to-face interactions) and non-physical interactions (respectively 10.3%), differences are more significant considering the frequency of interaction. Globally, non-physical interactions are more frequent than physical ones. More precisely, on the scale detailed in subsection 3.3, we observe that 135 partner couples (17.9%) interacted less than once a year in face-to-face and only 90 (11.9%) of them when considering interactions through the use of ICT. At the opposite, only 12.9 per cent of the pairs of partners interacted more than once a month in face-to-face, against 23.5 per cent for non-physical interactions. This result confirms the necessity to take into account the impact of geographical proximity on the frequency of these two types of interactions.

We want to determine the factors explaining this heterogeneity of interactions. Our object is to test econometrically the impact of different determinants on both the existence and frequency of interactions in a collaborative project. We therefore ran an ordered probit with a sample selection to identify which factors (primarily, measures of proximity) could explain first the existence of interactions between two partners, and second their frequency. The variable to be explained is discrete and ordered, and data observability is restricted by a binary selection

Table 5. Centrality indicators on theoretical and empirical graphs

| ID | Betweenness centrality | | | Degree centrality | | | Closeness centrality | | | | |
|------|------------------------------|-----------------|----------------|-------------------|------------------------------|-----------------|----------------------|------|------------------------------|-----------------|----------------|
| | Complete graph | High hyp. graph | Low hyp. graph | ID | Complete graph | High hyp. graph | Low hyp. graph | ID | Complete graph | High hyp. graph | Low hyp. graph |
| A267 | 28.55 (1) | 28.57 (1) | 24.01 (2) | A267 | 28.74 (1) | 28.74 (1) | 2.83 (2) | A267 | 51.58 (1) | 49.43 (1) | 47.52 (2) |
| A268 | 17.05 (2) | 17.56 (2) | 18.60 (4) | A49 | 20.31 (2) | 20.31 (2) | 2.87 (1) | A49 | 49.62 (2) | 48.42 (2) | 47.81 (1) |
| A264 | 15.90 (3) | 15.95 (4) | 18.48 (5) | A264 | 18.39 (3) | 17.62 (3) | 2.13 (3) | A31 | 48.24 (3) | 45.95 (3) | 46.42 (3) |
| A31 | 15.81 (4) | 13.13 (5) | 21.93 (3) | A31 | 16.86 (4) | 16.48 (4) | 1.92 (4) | A264 | 45.79 (4) | 44.69 (5) | 43.72 (4) |
| A49 | 15.53 (5) | 16.97 (3) | 27.84 (1) | A268 | 16.48 (5) | 15.33 (5) | 1.50 (6) | A226 | 45.31 (5) | 45 (4) | 41.67 (6) |
| A265 | 13.36 (6) | 8.67 (8) | 12.39 (6) | A265 | 16.48 (6) | 14.94 (6) | 1.46 (8) | A266 | 44.69 (6) | 44.31 (6) | 41.45 (7) |
| A266 | 10.08 (7) | 12.04 (6) | 9.98 (8) | A266 | 15.71 (7) | 14.56 (7) | 1.79 (5) | A268 | 44.24 (7) | 43.94 (7) | 42.25 (5) |
| A226 | 9.31 (8) | 10.46 (7) | 6.00 (12) | A89 | 12.64 (8) | 10.73 (10) | 1.46 (7) | A233 | 42.78 (8) | 41.89 (8) | 37.09 (22) |
| A80 | 8.10 (9) | 8.10 (10) | 8.78 (10) | A235 | 11.49 (9) | 11.11 (9) | 0.88 (13) | A240 | 41.89 (9) | 41.03 (9) | 37.8 (18) |
| A89 | 7.11 (10) | 8.44 (9) | 10.23 (7) | A226 | 11.49 (10) | 11.11 (8) | 0.92 (11) | A69 | 41.82 (10) | 39.48 (16) | 38.09 (15) |
| | Kendall's rank correlation | | | | Kendall's rank correlation | | | | Kendall's rank correlation | | |
| | Complete/High hyp.: 0.786*** | | | | Complete/High hyp.: 0.825*** | | | | Complete/High hyp.: 0.919*** | | |
| | Complete/Low hyp.: 0.716*** | | | | Complete/Low hyp.: 0.319*** | | | | Complete/Low hyp.: 0.772*** | | |

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 6. Distribution of interaction frequency by ties

| Scale of frequency | Physical interactions | | Non-physical interactions | |
|------------------------------|-----------------------|-------|---------------------------|-------|
| | n | % | n | % |
| 0: No interaction at all | 64 | 8.5 | 78 | 10.3 |
| 1: Very few interactions | 135 | 17.9 | 90 | 11.9 |
| 2: Few interactions | 253 | 33.5 | 158 | 21.0 |
| 3: Regular interactions | 205 | 27.2 | 251 | 33.3 |
| 4: Very regular interactions | 97 | 12.9 | 177 | 23.5 |
| Total | 754 | 100.0 | 754 | 100.0 |

mechanism (De Luca and Perotti 2010). The introduction of a selection equation allows the potential bias of the existence of interactions to be taken into account before studying their frequency.

An ordered response model with sample selection can be represented by the following bivariate threshold crossing model:

$$Y_j^* = \beta_j X_j + \mu_j \text{ with } j = 1, 2 \tag{1}$$

$$Y_1 = I(Y_1^* \geq 0), \tag{2}$$

$$Y_2 = \sum_{h=0}^H hI(\alpha_h < Y_j^* \leq \alpha_{h+1}) \text{ if } Y_1 = 1 \tag{3}$$

where Y_1^* and Y_2^* represent continuous latent variables for the selection process and the outcome of interest respectively, β_j are k_j vectors of unknown parameters, X_j are k_j vectors of exogenous variables, and μ_j represents random errors (Equation (1)). The latent variable Y_1^* is related to the binary indicator Y_1 through the observational rule (Equation (2)), and $I(A)$ denotes the indicator function of the event A. The latent variable Y_2^* is related to the outcome Y_2 through the observational rule (Equation (3)), where $\alpha = (\alpha_1, \dots, \alpha_H)$, with $\alpha_h < \alpha_{h+1}$, $\alpha_0 = -\infty$ and $\alpha_{H+1} = +\infty$ is a vector of H with strictly increasing thresholds which partitions Y_2^* into H+1 exhaustive and mutually exclusive intervals. As in a classical sample selection model, the observability of Y_2 is confined to the sub-sample of observations for which $Y_1 = 1$ (the selected sample). Selectivity effects operate *via* the correlation between the latent regression errors μ_1 and μ_2 .

The unit of analysis is the dyad of partners jointly participated in a project. In the selection Equation (2) Y_1 which concerns the existence of interactions between two partners, the explained variable takes the value of 1 if there are interactions and 0 if the partners do not interact. In the outcome Equation (3) Y_2 which concerns the frequency of such interactions, the explained variable takes the value 1, 2, 3 or 4 according to the frequency scale used in the survey. We use the same explanatory variables in the two equations, and test two versions of the model: the first one focuses on physical interactions while the second one focuses on non-physical interactions.

5.2 The determinants of frequency of interactions

Before commenting on the results presented in Table 7, we must note that the inclusion of a selection equation, legitimized by the nature of our dependent variable, does not introduce bias

Table 7. Estimation of ordered probit with sample selection

| | | Physical interactions | | Non-physical interactions | |
|-------------------------|-----------------|-------------------------|-------------------------|---------------------------|-------------------------|
| | | Equation (1) n = 754 | Equation (2) n = 690 | Equation (1) n = 754 | Equation (2) n = 676 |
| Project characteristics | project_size | | | | |
| | ≤5 | ref. | ref. | ref. | ref. |
| | >5 | 0.20 | 0.37*** | 0.18 | 0.18 |
| | funding | | | | |
| | cc_policy | ref. | ref. | ref. | ref. |
| | research_agency | 0.30 | -0.37** | 0.03 | -0.34** |
| | europe | -0.71*** | -0.55*** | -0.85*** | -0.34 |
| | local | -0.33 | 0.17 | -0.84** | -0.25 |
| | period_label | | | | |
| | period1 | ref. | ref. | ref. | ref. |
| period2 | 0.29 | 0.08 | 0.13 | 0.18 | |
| colabeling | -0.03 | -0.51*** | -0.14 | -0.29** | |
| Tie variables | coordinator | 0.90*** | 0.59*** | 0.98*** | 0.56*** |
| | geo_proxi | -0.06 | 0.33** | 0.06 | 0.05 |
| | partnership | | | | |
| | sc_sc | ref. | ref. | ref. | ref. |
| | ind_ind | -0.20 | -0.47*** | -0.13 | -0.13 |
| | sc_ind | 0.06 | -0.20 | -0.05 | -0.16 |
| | acquaintance | 0.41** | 0.47*** | 0.59*** | 0.45*** |
| constant | 0.74* | - | 0.85** | - | |
| | | Wald Chi2 = 135.59*** | | Wald Chi2 = 63.86*** | |
| | | Rho = 0.10 | | Rho = 0.23 | |

Notes: Equation (1): selection equation on existence; Equation (2): output equation on frequency. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(rho is not significant) for any of the two models. In other words, the results obtained in Equation (2) correspond perfectly with those obtained from the binary probit tested in subsection 4.1; an ordered probit without the selection equation would have given the same results as Equation (3). For each variable, the results differ depending on whether it has a significant impact on both the existence and frequency of interactions, or on only one of the two, or on neither.

As seen above, the project size does not impact on the existence of interaction, but it is the case for the frequency of physical interactions between partners: face-to-face interactions are significantly more frequent for the biggest projects than for the smallest ones. We can assume that more a project involves partners, more the co-ordination complex, and higher the need for face-to-face meetings.

With regard to the type of funding, we observe significant and negative impacts on the frequency of interactions – both on physical and non-physical ones – for projects funded by the French national research agency. The fact there are less frequent interactions than in CC policy projects can be explained by the fact that the CC’s interventions mean the projects are more structured. In fact, CCs give their projects a great deal of support and attention, as they act as a showcase for the cluster’s identity and success: the support by this bottom-up policy tends to reinforce the levels of interaction between partners. For European projects, we have already commented on their tendency to adopt sub-project structures, which may explain why some partner couples in the same project do not necessarily interact at all. Concerning the frequency of interactions, this variable has a differential effect on the two types of interactions: face-to-face interactions are significantly less frequent in European projects, but there is no impact on non-physical interactions.

Co-labelling – where project partners are members of both the CC we studied and of other CCs – has a negative impact on the frequency of interactions between partners, both on physical and non-physical ones. In 2009, the French government encouraged inter-clustering – collaborations between partners from different CCs – and this policy orientation appears to have had a real influence, as the proportion of co-labelled projects subsequently increased from 27 to 35 per cent. While such an increase could have been obtained by artificially linking some actors to build inter-cluster projects, it mainly represents new collaborations in which players first have to get to know each other: that may explain the lower levels of interaction in such projects than in non-co-labelled ones, in which partners are more likely to benefit from previous acquaintance.

The role of co-ordinator also seems to be important in projects: the frequency of interactions is significantly higher for ties involving co-ordinators, confirming previous results and supporting the need to combine the star graph and complete graph representations. This effect is important for both physical and non-physical interactions.

Finally, previous acquaintance between partners has the highest and positive coefficient in the two models, highlighting the importance of social relationships and mutual confidence in supporting co-ordination. This result corroborates Reinold et al. (2013) and Paier et al. (2011) on the role of previous collaboration on inter-organizational generation of scientific knowledge for the first study, and on the intensity of knowledge exchange for the second one.

The nature of the partnership (science and/or industry) affects neither the existence nor the frequency of non-physical interactions. In case of physical meetings, it seems that the different nature of partnership affects the frequency of interaction: industry-industry partners interact in face-to-face significantly less than science-science ones, confirming the results of Reinold et al. (2013).

Confirming the results of the large majority of studies on geographical proximity, Table 7 shows that geographical proximity has a positive effect on the frequency of face-to-face interactions, but no effect on non-physical interactions. Being co-located increases the likelihood that physical meetings occur between partners; but knowledge can also be exchanged through the media of ICT whatever the geographical distance between partners, with a minimum of face-to-face interactions. It relativizes the previous result concerning the non-significant impact of geographical proximity on existence of interactions: geographical proximity is not a necessary condition to interact and exchange knowledge, but it can favour the frequency (and so the intensity) of some interactions.

6 Conclusion

The main objective of this study is to gather empirical declarative data to better understand the nature of interactions in collaborative innovation projects and their determinants. Our research is based on a survey addressed to partners involved in projects labelled by a French competitiveness cluster, from which we gained information on the existence and frequency of interactions between partners in 754 collaborative ties. This case study is not intended to make judgments on the French national cluster policy, but rather to learn from original and current material from this source.

Regarding our results, we can consider that complete graph representations improperly assume on average 7 per cent of ties to be active. This amount may seem at the same time both negligible and decisive: negligible because it represents a small proportion of the total links, encouraging the validation of the complete graph representation; decisive because network properties can be easily disrupted by the deletion of only a few strategic ties. So we recommend being cautious about the use of complete graphs: even if our study demonstrates that the two

empirical graphs – as obtained from the actors' own declarations – are not very different from the complete graph, it confirms that interactions between partner couples are far from being homogeneous. We show that the hypothesis of the complete graph representation cannot be fully rejected when dealing with collaborative funded projects, but the use of this representation hides the heterogeneous nature of inter-organizational ties in terms of the frequency and the channel of interaction.

In terms of the determinants of interaction, we observe the following three variables have stable and significant impacts on both the existence and frequency of interactions. First, interactions are less likely to exist and are more infrequent in European projects than in other projects: nearly one quarter of ties within European funded projects are characterized by an absence of any interaction against only 7 per cent for the all sample. This result is especially important, given the huge literature focusing on European Framework Programmes. The precaution recommended previously about the application of complete graph hypothesis must be taken with more caution using European Framework data. Second, co-ordinators generally appear to have important structuring roles in projects: on average, partners interact more with them than they do with the other actors. While ties involving co-ordinators are not the only ones that exist, they are usually significantly stronger. This result legitimates the underlying idea of star graph, but – as we demonstrate – this representation risks underestimating the number of existing ties: superimposing the star graph on the complete graph would give a more accurate representation of the weight of network ties. Third, previous acquaintance between partners is the most significant determinant of the frequency of their interactions during collaborative projects. Having previous collaborative experience facilitates the operation of the current project, supporting arguments about the importance of sociological dimensions.

Concerning geographical proximity, this key variable does not impact the likelihood that two partners interact within a collaborative project. It does not play a significant role on the frequency of non-physical interactions but it increases the frequency of face-to-face interactions, confirming previous results on the literature about geography of innovation and suggesting geography still matters (Morgan 2004). This result also highlights the necessity to take into account the channel of interaction when studying geography of innovation processes.

In terms of policy implications, one of the main objectives of clusters policy is to expand their networks and gain new members, but our findings suggest they should also focus on consolidating existing ties based on social relationships, and stress the benefits of meetings organized within the CC framework which encourage actors to exchange in informal ways. Bottom-up policy in the French context (for example through the CC policy funding of projects) allows reinforcing existing ties and facilitating circulation of knowledge inside the cluster.

While this paper introduces methodological insights into the network analysis of collaborative innovation projects, several limitations need to be noted. We face the traditional disadvantage of survey research – the incompleteness of answers – so that our data does not cover the whole network of the studied CC. Comparison with other CCs, and controlling by sectorial specialization and location, would consolidate the findings of this research. Finally, we chose to use the frequency of interactions to define their intensity. This measure is more quantitative than qualitative and further research should study the relationship between the quantity and the quality of interactions. This issue is all the more important, given that the literature predominantly links collaborative ties with knowledge diffusion. This association can be extended to collaboration success, prompting consideration as to whether more frequent interactions make initial project ambitions more likely to be realized. It highlights the need to combine quantitative and qualitative methods when evaluating innovation policy.

References

- Aguilera A, Lethiais V (2011) Transmission des connaissances dans les relations de coopération inter-entreprises: TIC vs. face-à-face. *Revue d'Economie Régionale et Urbaine* 2: 269–293
- Ahuja G (2000) Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly* 45: 425–455
- Arvanitis S, Kubli U, Woerter M (2008) University-industry knowledge and technology transfer in Switzerland: what university scientists think about co-operation with private enterprises. *Research Policy* 37: 1865–1883
- Autant-Bernard C, Mairesse J, Massard N (2007a) Spatial knowledge diffusion through collaborative networks. *Papers in Regional Science* 86: 341–350
- Autant-Bernard C, Billand P, Frachisse D, Massard N (2007b) Social distance versus spatial distance in R&D co-operation: Empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in Regional Science* 86: 495–520
- Balland PA (2012) Proximity and the evolution of collaboration networks: Evidence from research and development projects within the global navigation sSatellite system (GNSS) industry. *Regional Studies* 46: 741–756
- Balland PA, Suire R, Vicente J (2013) Structural and geographical patterns of knowledge networks in emerging technological standards: Evidence from the European GNSS industry. *Economics of Innovation and New Technology* 22: 47–72
- Bekkers R, Freitas I (2008) Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter? *Research Policy* 37: 1837–1853
- Belderbos R, Carree M, Lokshin B (2004) R&D co-operation and firm performance. *Research Policy* 33: 1477–1492
- Boardman C, Bozeman B (2006) Implementing a ‘bottom-up’ multi-sector research collaboration: The case of the Texas air quality study. *Economics of Innovation and New Technology* 15: 51–69
- Borgatti SP, Everett MG, Freeman LC (2002) *Ucinet for Windows: Software for social network analysis*. Analytic Technologies, Harvard, MA
- Borgatti SP, Foster PC (2003) The network paradigm in organizational research: A review and typology. *Journal of Management* 29: 991–1013
- Boschma R (2005) Proximity and innovation: A critical assessment. *Regional Studies* 39: 61–74
- Boschma R, Frenken K (2010) The spatial evolution of innovation networks: A proximity perspective. In: Boschma R, Martin R (eds) *Handbook of evolutionary economic geography*. Edward Elgar, Cheltenham
- Bouba-Olga O, Ferru M, Pepin D (2012) Exploring spatial features of science-industry partnerships, a study on French data. *Papers in Regional Science* 91: 355–375
- Bouba-Olga O, Grossetti M (2008) Socio-économie de proximité. *Revue d'Economie Régionale et Urbaine* 3: 311–318
- Brenner T, Cantner U, Fornahl D, Fromhold-Eisebith M, Werker C (2011) Regional innovation systems, clusters, and knowledge networking. *Papers in Regional Science* 90: 243–249
- Breschi S, Cusmano L (2004) Unveiling the texture of a European research area: Emergence of oligarchic networks under EU framework programmes. *International Journal of Technology Management* 27: 747–772
- Breschi S, Lissoni F (2009) Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography* 9: 439–468
- Busom I, Fernandez-Ribas A (2008) The impact of firm participation in R&D programmes on R&D partnerships. *Research Policy* 37: 240–257
- Cairncross F (1997) *The death of distance: how the communications revolution will change our lives*. Harvard Business School Publishing, Cambridge, MA
- Carayol N, Roux P (2008) The strategic formation of inter-individual collaboration networks, evidence from co-invention patterns. *Annales d'Economie et de Statistique* 87: 275–301
- Cassiman B, Di Guardo MC, Valentini G (2010) Organizing links with science: Co-operate or contract? A project-level analysis. *Research Policy* 39: 882–892
- Cassiman B, Veugelers R (2002) R&D Co-operation and spillovers: Some empirical evidence from Belgium. *The American Economic Review* 92: 1169–1184
- Cunningham W, Werker C (2012) Proximity and collaboration in European nanotechnology. *Papers in Regional Science* 91: 723–742
- De Luca G, Perotti V (2010) Estimation of ordered response models with sample selection. Working Paper 168, CEIS
- Eisenhardt K (1989) Building theory from case study research. *Academic Management Review* 14: 532–550
- EuroLIO (2010) L’innovation dans les pôles de compétitivité. URL: <http://www.eurolio.eu/Productions/Fiche-Pole>
- Ferru M (2014) Partners’ connection process and spatial effects: New insights from a comparative inter-organizational partnerships analysis. *European Planning Studies* 22: 975–994
- Friedkin NE (1982) Information flow through strong and weak ties in intraorganizational social networks. *Social Networks* 3: 273–285

- Fritsch M, Kauffeld-Monz M (2010) The impact of network structure on knowledge transfer: An application of social network analysis in the context of regional innovation networks. *Annals of Regional Science* 44: 21–38
- Gay B, Dousset B (2005) Innovation and network structural dynamics: Study of the alliance network of a major sector of the biotechnology industry. *Research Policy* 34: 1457–1475
- Gilly JP, Torre A (2000) *Dynamiques de proximité*. L'Harmattan, Paris
- Gilsing V, Nootboom B, Vanhaverbeke W, Duysters G, Van den Oord A (2008) Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy* 37: 1717–1731
- Gomes-Casseres B, Hagedoorn J, Jaffe A (2005) Do alliances promote knowledge flows? *Journal of Financial Economics* 80: 5–33
- Grandclement A (2011) Vers une géographie des pôles de compétitivité : Réseaux, construction politique et organisation territoriale. In: Resmini L, Torre A (eds) *Competitività territoriale: determinanti e politiche*, Franco Angeli, Milan
- Granovetter M (1983) The strength of weak ties: A network theory revisited. *Sociological Theory* 1: 201–233
- Granovetter M (1985) Economic action and social structure: the problem of embeddedness. *American Journal of Sociology* 91: 533–544
- Grossetti M, Bès MP (2003) Dynamique des réseaux et des cercles: Encastrement et découplages. *Revue d'Economie Industrielle* 103: 43–58
- Gulati R (1995) Social structure and alliance formation patterns: A longitudinal analysis. *Administrative Science Quarterly* 40: 619–652
- Hagedoorn J (2006) Understanding the cross-level embeddedness of inter-firm partnership formation. *Academy of Management Review* 31: 670–680
- Hagedoorn J, Roijakkers N (2006) Inter-firm R&D partnering in pharmaceutical biotechnology since 1975: trends, patterns, and networks. *Research Policy* 35: 431–446
- Hanaki N, Nakajima R, Ogura Y (2010) The dynamics of R&D network in the IT industry. *Research Policy* 39: 386–399
- He ZL, Geng XS, Campbell-Hunt C (2009) Research collaboration and research output: A longitudinal study of 65 biomedical scientists in a New Zealand university. *Research Policy* 38: 306–317
- Hoekman J, Frenken K, Tijssen RJW (2010) Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy* 39: 662–673
- Hussler C, Ronde P (2007) Explaining the geography of co-patenting in the scientific community: A social network analysis. *Annales d'Economie et de Statistique* 87: 351–372
- Katz JS (1994) Geographical proximity and scientific collaboration. *Scientometrics* 31: 31–43
- Kirat T, Lung Y (1999) Innovation and proximity, territories as loci of collective learning processes. *European Urban and Regional Studies* 6: 27–38
- Kline S, Rosenberg N (1986) An overview of innovation. In: Landau R, Rosenberg N (eds) *The positive sum strategy: Harnessing technology for economic growth*. National Academy Press, Washington, DC
- Levy R (2005) Les doctorants CIFRE: Médiateurs entre laboratoires de recherche universitaires et entreprises. *Revue d'Economie Industrielle* 111: 79–96
- Levy R, Roux P, Wolff S (2009) Study of science-industry collaborative patterns in a large European university. *Journal of Technology Transfer* 34: 1–23
- Levy R, Talbot D (2014) Control by proximity: Evidence from the 'Aerospace Valley' competitiveness cluster. *Regional Studies*. doi:10.1080/00343404.2013.840721
- Maggioni M, Uberti T (2009) Knowledge networks across Europe: Which distance matters? *The Annals of Regional Science* 43: 691–720
- Mora-Valentin EM, Montoro-Sanchez A, Guerras-Martin LA (2004) Determining factors in the success of R&D co-operative agreements between firms and research organizations. *Research Policy* 33: 17–40
- Morgan K (2004) The exaggerated death of geography: Learning, proximity and territorial innovation systems. *Journal of Economic Geography* 4: 3–21
- OECD (1996) *The knowledge-based economy*. OECD, Paris
- Paier M, Scherngell T (2011) Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model. *Industry and Innovation* 18: 89–104
- Ponds R, Van Oort F, Frenken K (2007) The geographical and institutional proximity of research collaboration. *Papers in Regional Science* 86: 423–443
- Porter M (1998) Clusters and the new economics of competition. *Harvard Business Review* 76: 77–90
- Reinold F, Paier M, Fischer MM (2013) Joint knowledge production in European R&D networks: Results from a discrete choice modeling perspective. In: Scherngell T (ed) *The geography of networks and R&D collaborations*. Springer International Publishing, Switzerland, pp 201–220
- Rodriguez-Pose A, Crescenzi R (2008) Research and Development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Regional Studies* 42: 51–67

- Roediger-Schluga T, Barber MJ (2008) R&D collaboration networks in the European Framework Programmes: Data processing, network construction and selected results. *International Journal of Foresight and Innovation Policy* 4: 321–347
- Schartinger D, Rammer C, Fischer MM, Frohlich J (2002) Knowledge interactions between universities and industry in Austria: Sectoral patterns and determinants. *Research Policy* 31: 303–328
- Scherngell T, Barber MJ (2009) Spatial interaction modelling of cross-region R&D collaborations: Empirical evidence from the 5th EU framework programme. *Papers in Regional Science* 88: 531–546
- Stuart TE, Ozdemir SZ, Ding W (2007) Vertical alliance networks: The case of university-biotechnology-pharmaceutical alliance chains. *Research Policy* 36: 477–498
- Ter Wal AL, Boschma R (2009) Applying social network analysis in economic geography: Framing some key analytic issues. *The Annals of Regional Science* 43: 739–756
- Thune T (2007) University-industry collaboration: The network embeddedness approach. *Science and Public Policy* 34: 158–168
- Torre A (2008) On the role played by temporary geographical proximity in knowledge transmission. *Regional Studies* 42: 869–889
- Torre A (2011) The role of proximity during long-distance collaborative projects: Temporary geographical proximity helps. *International Journal Foresight and Innovation Policy* 7: 213–230
- Torre A, Rallet A (2005) Proximity and localization. *Regional Studies* 39: 47–60
- Vicente J, Balland PA, Brossard O (2011) Getting into networks and clusters: Evidence from the Midi-Pyrenean global navigation satellite systems (GNSS) collaboration network. *Regional Studies* 45: 1059–1078
- Vonortas NS (2013) Social networks in R&D program evaluation. *Journal of Technology Transfer* 38: 577–606
- Wagner CS, Leydesdorff L (2005) Network structure, self-organisation, and the growth of international collaboration in science. *Research Policy* 34: 1608–1618
- Yin RK (2003) *Case study research, design and methods*. Sage, Thousand Oaks, CA



Resumen. Se discute la hipótesis común de una representación gráfica completa, según la cual, en proyectos de colaboración, todos los socios interactúan entre sí de forma homogénea. Más específicamente, el objetivo de este estudio es determinar la heterogeneidad en cuanto a la existencia y la frecuencia de interacciones entre las díadas de organizaciones que participaron en forma conjunta en proyectos de colaboración. A partir de una encuesta a participantes en proyectos de innovación aprobados por un conglomerado francés, se recopiló información sobre 754 vínculos de colaboración. Mediante el uso de un probit ordenado con ecuación de selección, se analizó a continuación el impacto de varios factores determinantes, entre ellos la proximidad geográfica, sobre la existencia y la frecuencia de las interacciones observadas.

要約: 本論文では、協調的プロジェクトにおいて、すべてのパートナーが同質的に相互作用する際に従う完全グラフの表現に関する一般仮説を考察する。正確には、本研究の目的は、協調的プロジェクトに共同参画する組織の1組のペアの相互作用の存在と頻度の観点での異質性を決定することである。フランスのクラスターによって承認されたイノベーションプロジェクトの参加者を対象にした調査から、754組のコラボレーションに関する情報を収集した。選択方程式 (selection equation) のある順序型プロビットを用いて、地域的な近接性などの要因が、調査対象の相互作用の存在と頻度に与える影響を検定する。