Exploring the dynamics of regional R&D networks: a closer look at Valencia's inter-organizational partnerships

Dima Nedelcheva Yankova

Author

Pablo D´Este & Mónica García Melón

Supervisors

Ph.D Dissertation

Programa de Doctorado en Desarrollo Local y Cooperación Internacional

Universitat Politècnica de València

Valencia, October 2023

Acknowledgements

I was fortunate enough to pursue my PhD as part of an outstanding Marie Curie Innovative Training Network, but unfortunate to commence amidst a global pandemic which added unforeseen layers of mental strain. I, therefore, take this opportunity to thank all those people who helped me navigate through the challenges of the last three years.

To Pablo: for his words of encouragement, unwavering patience and meticulous attention to detail. I owe you a massive debt of gratitude for teaching me so much about the *craft* of scientific research. To Mónica, *my Spanish mom*: for her steadfast support and nurturing care for my professional career during and after the PhD. Thank you both for believing in me since that very first interview and granting me the freedom to explore, stumble, learn and grow.

To my POLISS fellows: it has been a privilege to journey down this path alongside all of you in hectic conferences, exhausting trainings, lengthy summer and winter schools. I have and continue to learn so much from you. A special note of gratitude goes to Ghinwa, the only person who was physically present to witness the countless ups and downs which marked my PhD experience, and always reminded me to persevere.

To everyone in and outside INGENIO who challenged my ideas, reviewed my papers, offered words of advice or criticism: thank you for helping me improve my work. I am also grateful to Kati, who welcomed me during my academic secondment at the University of Pécs and provided me with the space to develop my personal and professional network.

And finally, to my family, which I've seen almost exclusively through the screen of a phone for the past 13 years, while pursuing my education abroad: this achievement is as much yours as it is mine.

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European's Union H2020 research and innovation programme under grant agreement No. 860887

POLISS POLICIES FOR SMART SPECIALISATION

Table of contents

List of figures

List of tables

Summary

The literature on knowledge networks has long grappled with two types of questions. The first concerns the antecedents of tie formation; that is, how actors select their partners. The second concentrates rather on the implications of the resulting network structure for knowledge exchange and individual or collective performance. Many studies have acknowledged the critical role of social proximity as a driver of link formation and an important prerequisite for the transfer of both tacit and complex knowledge. Yet, scholarly understanding of social proximity as a concept remains somewhat constrained and the implications of building strong ties are subject to ongoing debates.

Hence, the primary objective of this doctoral work is to address two sets of research questions. First, we aim to investigate how various forms of social proximity influence the formation of ties within knowledge networks. In this context, we differentiate between prior joint experiences in successful and unsuccessful project applications, as both forms of engagement constitute a source of relational embeddedness between actors. Second, we examine how the emerging strong bonds between organizations differ in their role and function. We test whether and under what conditions organizations leverage repeated collaborations to exploit the same topic multiple times (what we call *specialization*) or to explore new ones (*diversification*). These questions contribute to two separate streams of literature: the one on knowledge network dynamics by highlighting the origin and consequences of strong coupling; and the one on strategic management by tracing organizations' strategic response to funding rejection.

The thesis zooms in on Valencia's regional publicly-funded R&D network. To conduct the empirical analysis, we build a unique dataset which contains information on all R&D partnerships, formed between 2016 and 2022, which requested public subsidy from one of the top two regional sources of innovation-related funding. The two entities together manage 75% of the 1.6 billion Euros designated for the implementation of the regional smart specialization strategy.

Overall, this document introduces a new, vastly unexplored facet of social proximity, thus challenging existing assumptions on what type of former interaction is necessary to generate sufficient levels of trust and familiarity so as to motivate further engagement between actors. Moreover, it demonstrates empirically that structurally equivalent network ties can assume fundamentally distinct roles, leading either to thematic specialization or diversification. These findings suggest that the danger of over-embeddedness in one type of activity after several collaborations may not necessarily be a product of the structural setting alone and the presence of strong ties. It is rather a product of organizations' strategic choices about how they harness their strong bonds. The conclusions of this thesis hold far-reaching implications for policy design, and can guide policymakers in steering more effective network interventions.

Keywords: knowledge network; R&D collaboration; social proximity; tie strength; repeated collaboration

Resumen

La literatura sobre redes de conocimiento se ha centrado históricamente en dos líneas de investigación. La primera analiza los orígenes de las conexiones entre nodos, explorando cómo los actores eligen a sus socios. La segunda se centra en las implicaciones de la estructura de red resultante para el intercambio de conocimiento y el rendimiento individual o colectivo. Muchos estudios han reconocido el papel fundamental que juega la proximidad social como impulsora de la creación de conexiones y como requisito para la transferencia de conocimiento tanto tácito como complejo. Sin embargo, el entendimiento académico de la proximidad social como concepto y las implicaciones de establecer vínculos fuertes sigue siendo relativamente limitado.

El objetivo principal de esta tesis es abordar dos preguntas de investigación. En primer lugar, pretendemos investigar el impacto de diversas formas de proximidad social en la formación de vínculos dentro de las redes de conocimiento. En este contexto, distinguimos entre las experiencias compartidas en solicitudes de proyectos exitosas y no exitosas, ya que ambas interacciones constituyen fuentes de conexión social entre los participantes. En segundo lugar, analizamos las diferencias entre las conexiones fuertes respecto a su papel y función. Investigamos si, y bajo qué condiciones, las organizaciones aprovechan las colaboraciones repetidas para explotar varias veces el mismo tema (lo que denominamos *especialización*) o para explorar nuevas áreas temáticas (*diversificación*). Estos aspectos contribuyen a dos corrientes importantes de la literatura: la dinámica de las redes de conocimiento, al destacar el origen y las implicaciones de los vínculos fuertes; y la gestión estratégica, al analizar las respuestas estratégicas de las organizaciones frente al rechazo de financiamiento.

La tesis se enfoca en la red regional valenciana de I+D financiada con fondos públicos. Para llevar a cabo el análisis empírico, creamos una base de datos que contiene información sobre todas las asociaciones de I+D formadas entre 2016 y 2022. Estas asociaciones solicitaron subvenciones públicas a dos entidades regionales, que administran conjuntamente el 75% de los 1.6 millones de euros asignados para la implementación de la estrategia regional de especialización inteligente.

En general, esta investigación introduce un aspecto inexplorado de la proximidad social, que desafía las hipótesis existentes sobre el tipo de interacción previa necesaria para establecer un nivel adecuado de confianza y familiaridad que motive futuras colaboraciones entre los actores. Además, este estudio demuestra de manera empírica que los vínculos de red que son estructuralmente equivalentes pueden desempeñar roles fundamentalmente diferentes, dando lugar a la especialización o diversificación temática. Estos resultados indican que el peligro de una excesiva concentración en un tipo de actividad tras varias colaboraciones no es simplemente un resultado del entorno estructural o de la presencia de vínculos fuertes, sino que es el resultado de las decisiones estratégicas tomadas por las organizaciones sobre cómo aprovechar mejor sus vínculos fuertes. Las conclusiones de esta tesis tienen implicaciones significativas para el diseño de políticas y pueden orientar a los responsables políticos a dirigir intervenciones más efectivas sobre redes de conocimiento.

Palabras claves: red de conocimiento; colaboración en I+D; proximidad social; vínculos fuertes; colaboración repetida

Resum

La literatura sobre xarxes de coneixement s'ha centrat històricament en dues línies d'investigació. La primera analitza els orígens de les conexions entre nodes, explorant com els actors trien els seus socis. La segona es centra en les implicacions de l'estructura de xarxa resultant per a l'intercanvi de coneixements i el rendiment individual o col·lectiu. Molts estudis han reconegut el paper fonamental que juga la proximitat social com a impulsora de la creació de connexions i requisit per a la transferència de coneixement tant tàcit com complex. No obstant això, coneixement acadèmic de la proximitat social com a concepte i les implicacions d'establir vincles forts continuen sent relativament limitats.

L'objectiu principal d'aquesta tesi és abordar dues preguntes d'investigació. En primer lloc, pretenem investigar l'impacte de diverses formes de proximitat social en la formació de vincles dins les xarxes de coneixement. En aquest context, diferenciem entre les experiències compartides en sol·licituds de projectes exitosos i no exitosos, ja que ambdues interaccions constitueixen fonts de connexió social entre els participants. En segon lloc, analitzem les diferències entre els vincles forts segons el seu paper i funció. Investiguem si, i sota quines condicions, les organitzacions aprofiten les col·laboracions repetides per explotar diverses vegades el mateix tema (el que anomenem especialització) o per explorar noves àrees temàtiques (diversificació). Aquests aspectes contribueixen a dues corrents importants en la literatura: la dinàmica de les xarxes de coneixement, en destacar l'origen i les implicacions dels llaços forts; i la gestió estratègica, en analitzar les respostes estratègiques de les organitzacions davant el rebuig de finançament.

La tesi es centra en la xarxa regional valenciana de I+D finançada amb fons públics. Per dur a terme l'anàlisi empíric, s´ha creat una base de dades que conté informació sobre totes les associacions de I+D formades entre 2016 i 2022. Aquestes associacions van sol·licitar subvencions públiques de dues entitats regionals, que administren conjuntament el 75% dels 1,6 milions d'euros assignats per a la implementació de l'estratègia regional d'especialització intel·ligent.

En general, aquest document introdueix un aspecte inexplorat de la proximitat social, que qüestiona les hipòtesis existents sobre el tipus d'interacció prèvia necessària per establir un nivell adequat de confiança i familiaritat que motivi futures col·laboracions entre els actors. A més, aquest estudi demostra de manera empírica que els vincles de la xarxa que són estructuralment equivalents poden assumir rols fonamentalment diferents, conduint a l'especialització o diversificació temàtica. Aquests resultats indiquen que el perill d'una excessiva concentració en un tipus d'activitat després de diverses col·laboracions no és simplement un resultat de l'entorn estructural o la presència de vincles forts. En canvi, és el resultat de les decisions estratègiques preses per les organitzacions sobre com aprofitar els seus vincles forts. Les conclusions d'aquesta tesi tenen implicacions significatives per al disseny de polítiques i poden orientar els responsables polítics a l'hora de dirigir intervencions més efectives sobre xarxes de coneixement.

Paraules clau: xarxa de coneixement; col·laboració en I+D; proximitat social; vincles forts; col·laboració repetida

Chapter 1. General introduction

1.1 Preamble

Collaboration in knowledge production has seen an unprecedented rise in the last several decades (Fortunato et al., 2018; Wuchty et al., 2007). Innovation is increasingly understood as a collective process, where actors of different institutional and organizational background pool their skills, money and resources to bring novel products and services to the market. Collaborative research and development (R&D) offers a number of benefits, from achieving economies of scale and scope to lowering associated risks and costs (Martínez-Noya & Narula, 2018). Moreover, the complexity of today's economic, environmental and societal challenges requires an integrated approach which counts on input from universities, firms, governments and civil sector organizations.

This increase in research collaboration has been recognized and supported by policymakers at all levels of government. A classic example is Europe's Framework Program sequence, with Horizon Europe being the most recent addition boasting a record-breaking 95 billion EUR budget (European Commission, 2021). Other initiatives, like the Smart Specialization policies and the Partnerships for Regional Innovation (PRI) are also worth mentioning as a case of policy intervention aimed at developing R&D collaborations at both intra and inter-regional level (Foray et al., 2009; Pontikakis et al., 2022).

With the rise of collaborative R&D, scholarly interest in knowledge networks has also proliferated. Researchers from across disciplines are trying to understand not only how specific network structures come into being, but also what those structures mean for knowledge transfer processes and actor performance.

Understanding the antecedents of tie formation is fundamental for optimizing network design, ensuring efficient resource allocation amongst network members and identifying bottlenecks that may hinder the flow of resources. At the same time, emerging network structures do not have a uniform effect on innovative output and performance. Some type of links or network configurations between actors may be more conducive to knowledge transfer than others. Thus, for instance, strong ties between network members, often built as a result of repetitive engagement, have been shown to favor the exchange of tacit and complex knowledge (Becerra et al., 2008; Coleman, 1988; Dyer & Singh, 2011; Fritsch & Kauffeld-Monz, 2009). Yet, in some situations they can also block the inflow of new ideas and perspectives, hindering innovative performance (Gargiulo & Benassi, 1999; Uzzi, 1996). What mechanisms lead to the consolidation of strong ties, and how the latter affect network performance is still very much an open debate.

This doctoral work will examine the case of Valencia's regional publicly-funded R&D network. On the one hand, we look at how different forms of social proximity influence future tie formation. We differentiate between prior joint experience in successful and unsuccessful project applications as both types of engagement constitute a source of relational embeddedness between actors. On the other hand, we also look at how the emerging strong ties between organizations in the network differ in their role and function. We test whether organizations leverage repeated collaborations to exploit the same topic multiple times (what we call specialization) or to explore new ones (diversification)? Both types of questions help expand scholarly understanding on the emergence of strong coupling between network actors and the implications of nurturing repeated collaborations.

Since the empirical analysis is grounded in an existing policy-induced R&D network, the results of the thesis provide practical insights for policymakers. They highlight previously unobserved patterns in organizational behavior, including actors' response to rejection of government funding and their approach to re-engaging with prior partners. Understanding these micro-mechanisms of network dynamics can help policymakers steer action in desired directions and avoid situations of stagnation.

In this chapter we will outline the main stands of literature used to inform our study. We define the principal research question, as well as several sub-questions, which are analyzed in greater detail throughout the rest of the document. A brief introduction of methodological approaches is also included. The chapter concludes with a graphical representation of the thesis structure.

1.2 Conceptual framework

In the literature on collaborative knowledge networks, two questions have prompted extensive research. The first concerns the antecedents of link formation: how do actors select their partners? (1.2.1.) The second concentrates rather on the implications of the resulting network structure: how does it impact knowledge transfer, individual or collective performance? (1.2.2.) (Boschma & Frenken, 2010). We can argue that the first line of research tries to unpack how a network structure emerges from the input perspective, while the latter scrutinizes the significance of that structure for knowledge transfer and actor performance on the output side.

In this section we present the main theoretical developments in each of the two research lines and we outline the contributions we make to each one.

1.2.1 The antecedents of tie formation

The current literature on knowledge networks is particularly interested in their structure and the mechanisms driving tie formation. Two separate theoretical frameworks have emerged. One borrows from the proximity literature, while the second categorizes the antecedents from a network theory perspective. Although the two frameworks originate from different streams of research, they seem to share certain commonalities, as we explain below.

Proximity framework

Following Boschma (2005), this framework presents various forms of proximity – geographical, social, cognitive, institutional and organizational – as potential driving factors behind tie formation. In brief, geographical proximity indicates the physical distance between two network members; social proximity typically describes their familiarity with each other and common past experience; cognitive proximity refers to shared knowledge bases; institutional, the extent to which two actors operate under the same institutions, and organizational proximity reflects the degree to which actors are subjected to common hierarchical control.

Proponents of this framework argue that some optimal level of proximity is necessary to stimulate collaboration and knowledge transfer between disparate members. The emerging structure of networks can be, therefore, best described by the interplay of various dimensions. Consequently, many empirical studies have tried to independently operationalize and assess the role of different proximities for network structural evolution and qualify their nature as either substitutes or complements (Balland, 2012; Balland et al., 2013; Broekel & Boschma, 2012; Cantner et al., 2017; D'Este et al., 2013; Tsouri, 2022)

Network framework

The second framework approaches the question of tie formation from a network theory perspective, by categorizing three levels of factors: node, dyad and network (Broekel & Hartog, 2013). Node-level determinants refer to specific attributes of the individual node, which may affect its propensity to establish ties, such as size, absorptive capacity, institutional characteristics, strategic orientation, inherited capabilities, prior collaborative or industrial experience (Balland et al., 2016; Boschma & Ter Wal, 2007; Giuliani & Bell, 2005; Juhász, 2021; Owen-Smith & Powell, 2004).

Factors at the dyad level concern specific properties of the relationship between two nodes, rather than their individual traits. It revolves primarily around the similarity of two nodes' attributes – what social network scientists call *homophily*, although it could also be viewed as proximity. Actors that share the same trait (location, status, cognitive capacity or other) may gravitate toward each other. Hence, on a conceptual level, dyadic factors resemble very closely the proximity dimensions described by Boschma (2005).

Finally, the formation of ties in a knowledge network could also be influenced by structural factors stemming from specific network configurations. Such determinants are not associated with node attributes, neither are they derived from shared dyadic characteristics, but are contingent on the presence of a particular network configuration, such as transitivity or triadic closure (Ter Wal, 2014). This concept captures the tendency of two unconnected nodes with a mutual acquaintance to eventually establish a link with one another, resulting in triadic closure. The implication of this mechanism on the whole network is the emergence of denser "cliques" of highly interconnected nodes (clustering), where partners of partners become partners.

1.2.2 Contribution to theory: on the role of social proximity

A common thread across both frameworks is the recognition of social proximity as a key determinant of tie formation in collaborative knowledge networks. Social proximity, sometimes rebranded as social or relational embeddedness in the network theory framework, is a dyadic factor which denotes familiarity and trust between two organizations, based on friendship, kinship or past experience (Boschma, 2005). In other words, if two actors have a shared experience, they are more likely to engage in a future collaboration as opposed to two actors that have never interacted with each other before.

Both frameworks associate the presence of a prior social link between two actors with greater predictability, and lower risk of conflict and opportunism (Bstieler et al., 2017; R. Gulati, 1995; Usai et al., 2017; Uzzi, 1996). In addition, familiarity with the organizational routines of the partner entity can contribute to a smoother and more efficient collaboration.

So far, the notion of social proximity has been conceptualized and operationalized mostly in one dimension: prior partnership. This means that the social link between two actors is typically equated to a former collaboration. However, it is important to recognize that two actors can have a wide range of shared past experiences, equally likely to nurture trust and familiarity — the two principal components of social proximity. This leads us to question whether it is possible that unobserved forms of prior engagement between actors – different from a full-fledged collaboration, can dictate their partner selection behavior in significant ways? And if so, how does the effect compare to that of already established drivers of tie formation?

To enrich the discussion above, this doctoral thesis introduces a new, vastly unexplored, facet of relational embeddedness, based on prior joint experience in failed project applications. We use the term "failed" to refer to partnerships which were not awarded government funding for the execution of their project proposal. In this way we challenge existing assumptions on what type of former interaction is necessary to generate sufficient levels of trust and familiarity so as to spur further engagement between individual actors. Moreover, we argue that failure is a common phenomenon in collaborative research and more often than not, partners face rejection when competing for external financial support. Yet, we know very little about how failed applications shape actors' collaborative behavior and the network they belong to as a whole.

The thesis provides further insight to extant literature by explicitly comparing the effect of failed partnerships to that of successful ones. We juxtapose a neglected aspect of relational embeddedness (shared failure) to the more traditional form of social proximity, based on prior collaboration. Exploring the interaction between these two types of engagement is particularly important, because relying on successful partnerships alone may produce an incomplete image of the network´s structural dynamics. In the absence of information on failed applications, researchers may end up attributing some portion of newly formed ties to chance, when they could in fact be driven by undetected instances of unsuccessful prior partnerships. This would suggest that the role of relational embeddedness – in all its forms – is perhaps underrated. Alternatively, if failed partnerships negatively affect future tie formation, then we may conclude that relational embeddedness does not have a uniform effect and the type of past experience matters.

In sum, this doctoral work aims to incorporate an additional perspective on what constitutes social proximity, and thus improve our understanding on the real mechanisms driving tie formation in knowledge networks.

1.2.3 The implications of network structural properties for performance

The second key stream of research in the literature on collaborative knowledge networks concerns the impact of network structural properties for individual and collective performance. Here, two distinct approaches can be identified. The first more dominant one seeks to establish a connection between observed network structural properties and a desired outcome, be it innovative activity, overall performance or knowledge transfer. The second approach aims to understand the precise mechanisms through which structural properties produce a certain outcome and it concentrates rather on the nature and content of established ties[1].

Structuralist perspective

The so-called structuralist perspective dominates knowledge network studies. It emphasizes the study of network properties such as density, centrality, clustering, and hierarchy. Hence, it tends to infer knowledge transfer from the association between network structure – including tie strength – and performance. For instance, cohesive networks with many closed triads boost trust between members and generally facilitate the transfer of both tacit and complex knowledge (Becerra et al., 2008; Coleman, 1988). At the same time, open triads, also known as structural holes, have been seen as particularly conducive to the generation of original or novel ideas (Burt, 2004).

Connectionist perspective

The "connectionist" perspective looks beyond the structural or topological properties of the network, and treats ties as conduits or pipes through which knowledge and resources flow (Owen-Smith & Powell, 2004; Podolny, 2001; Snijders, 1999). This approach acknowledges that seemingly identical types of relationships or links could exercise distinct functions and transmit varying kinds of resources, affecting innovation outcome. Furthermore, it recognizes that organizations in alliance networks are not simply "helpless targets of structural influence", but active agents which make conscious decisions about the way they leverage strong bonds (Madhavan & Prescott, 2017).

1.2.4 Contribution to theory: strong ties and thematic specialization

The structuralist framework has been instrumental for establishing a clear relationship between various network configurations and the production of original or innovative output. However, it suffers from a number of limitations. Most notably, structuralists treat inter-organizational ties as virtually homogeneous, without classifying their individual functions within the network (Borgatti & Foster, 2003; Madhavan & Prescott, 2017; Phelps et al., 2012). Many studies tend to infer knowledge transfer from the association between network structure – including tie strength – and performance, without directly examining

[1] Despite the outlined distinctions between the two theoretical perspectives, it is worth mentioning recent attempts to integrate both views, particularly in the study of multiplex networks, whereby heterogeneity in tie content is acknowledged and interpreted as a driver of network evolution (see for instance Ferriani et al., 2012; Hammoud & Kramer, 2020).

the essence of this exchange, the interaction processes or the strategic decisions partners make (Ahuja et al., 2012; Madhavan & Prescott, 2017). One of the main premises of this dissertation is that the precise mechanisms underlying these associations can be fully understood only by bringing on board the connectionist perspective.

In this doctoral work we attempt to examine not just the presence of a link between two network members, but to further qualify the nature of that link. We do so by extracting the topic and content of partners' collaboration as a way of inferring the essence of their exchange. Thus, we can better understand how individual network members leverage their ties and what function these ties perform for the network as a whole. More specifically, we explore whether and under what conditions actors use their relational embeddedness (strong ties) to explore new topics (what we call diversification), as opposed to deepening their expertise in a single one.

Doing so contributes directly to the debate on the role of strong ties in inter-organizational networks. In the current academic discourse, strong bonds are linked to performance in an inverted U-shape. This means that organizations benefit from consolidating strong relationships up to a certain level, beyond which social embeddedness can act as a filter for the entry of new knowledge and perspectives, causing cognitive isolation and suboptimal innovative performance (Gargiulo & Benassi, 2000; Masciarelli et al., 2010). This situation is also known as "the proximity paradox", since the same factors that drive actors to connect and exchange knowledge may also lead them to innovate less in the long run (Boschma & Frenken, 2010; Broekel & Boschma, 2012).

As we argue further in this document, the relationship between strong ties and performance is not independent from the nature of the exchange between network members. Rather, it is contingent on it. Thus, for instance, if two actors systematically exploit the same topic and tap into the same knowledge domain, their interactions may be expected to yield benefits at first, but if continued for too long will likely hinder long-term innovative performance. If, on the other hand, subsequent collaborations begin to explore different topics, either because a priori the organizations involved possess a diverse internal repository of competencies and skills, or because they are capable of continuously sourcing novel knowledge through additional partnerships, the prospects of decreasing marginal benefits may weaken.

To summarize, this thesis examines both the role of social proximity as a driving factor in tie formation and the implications of the resulting strong bonds for knowledge transfer, and more specifically thematic specialization. Thus, it contributes to two separate strands of literature: one on the input side on the structure and evolution of knowledge networks and one on the output side, regarding the consequences of repeated ties.

1.3 Research questions

The goal of this doctoral thesis can be summarized in the following two connected research questions:

(1) How do different types of social proximity influence tie formation in knowledge networks and (2) what are the consequences of repeated collaborations for knowledge exchange processes?

The intention is, therefore, to enrich scholarly understanding on the mechanisms of tie formation in knowledge networks and to highlight the consequences of strong ties for knowledge exchange. To address the question at hand, two separate research lines have been developed, each containing a set of sub-questions.

Research line 1

(a) How does joint experience in failed project applications influence future tie formation?

(b) How does this relationship play out for partners with different levels of cognitive proximity?

Research line 2

(a) To what extent do organizations leverage repeated collaborations to exploit the same topic multiple times (specialization) or to explore new ones (diversification)?

(b) To what extent does partners' range of unique connections to other organizations inspire diversification in their repeated collaborations?

The first set of questions, 1(a) and 1(b), contribute to the *input* side of network analysis, by emphasizing how various forms of social proximity can interact to alter actors' collaborative behavior and impact the overall structural dynamics of the network. In contrast, the second set of questions 2(a) and 2(b) shift our focus to the *output* side. They investigate how the existing structure of the network, specifically the presence of strong ties, may affect knowledge transfer processes and lead to thematic specialization or diversification.

1.4 Research design and methodological approach

In this section we describe how the research questions outlined above have been addressed, and justify (1.4.1.) the empirical and (1.4.2.) methodological approaches that underlie the research design of the thesis.

1.4.1 Empirical setting

In extant literature, knowledge networks are generally constructed based on co-patenting, co-publishing or inter-organizational collaboration. In this doctoral work, we rely on the latter approach and we analyze a regional inter-organizational R&D network. So far, most studies of collaborative R&D networks have mobilized data from the European Framework Programs for Research and Technological Development (EU-FPs), taking countries as the unit of analysis, or considering inter- rather than intra-regional knowledge transfer (Autant-Bernard et al., 2007; Balland et al., 2019; Hoekman et al., 2010; Paier & Scherngell, 2011; Scherngell & Barber, 2011). Regional studies remain scarce, in part because longitudinal information on sub-national subsidized R&D projects is generally harder to find and is rarely organized as consistently or systematically.

The empirical analysis contained in this document relies on a single dataset. The set was built manually by extracting information on all R&D consortia, which applied for public funding to one of the two main sources of government subsidies in the Spanish region of Valencia: the Valencian Institute for Business Competitiveness (IVACE) and the Valencian Innovation Agency (AVI). Information was gathered directly from the online records of the two public organizations. The data covers a 7-year time window from 2016 to 2022. All collaborative partnerships were formed in response to government-sponsored programs, whose goal is to encourage downstream cooperation and boost regional innovation. Approximately half of the partnerships received public funding for the execution of their proposed R&D project, while the other half were not awarded a subsidy.

Chapter 1 provides a detailed overview of the dataset, and the resulting intra-regional R&D network. Chapter 2 exploits the dichotomy of "approved" vs "rejected" project applications, which is a distinctive feature of our data. Meanwhile, chapter 3 mobilizes only the information on "approved" projects, but brings on board textual evidence on the topic and content of each realized collaboration.

1.4.2 Methodological approach

The specificities of network data

Working with relational (network) data poses a unique set of challenges, which often require data reformatting and a careful selection of appropriate statistical methods. One such example of data conversion is the transformation of two-mode or bipartite network data (i.e. project-organization) into a one-mode projection with pairs of organizations. Although both approaches offer their own set of benefits, the use of one-mode projections can simplify the analysis by focusing on the relationship between organizations while preserving the relevant information from the original two-mode dataset. It is particularly useful for exploring collaborative patterns and interactions among organizations in the context of shared projects without the complexity of two-mode data.

Furthermore, relational data often suffers from structural interdependences. In dyadic datasets, autocorrelation between observations is rather common, since ties between nodes are often formed on the basis of existing ties. Although the use of standard inference tools in such cases remains commonplace, it is sometimes beneficial to supplement the empirical analysis with a statistical approach which directly accounts for the observed covariance. Broekel et al. (2014), for example, offer an overview of four distinct methods that have gained prevalence in the economic geography field. In this thesis we employ one of them, the Quadratic Assignment Procedure, to corroborate the results of a standard logistic regression.

Methods employed

Since our goal is to study the dynamics of a regional R&D network, the methods employed have been carefully selected to handle relational data. Depending on the objectives of each chapter, a mixture of inductive and deductive reasoning has been applied.

In chapter 2, for example, we use social network analysis (SNA) to uncover the overall structure of interactions between regional partners. We first map the relations between organizations and then employ an *inductive* approach to discover emerging patterns of behavior that could help illuminate the specific characteristics of the Valencian R&D network.

In chapters 3 and 4 we rely on a purely *deductive* approach. We formulate a set of hypotheses, which are firmly grounded in existing theory, and we adopt different econometric techniques to inspect the relationship between our variables of interest. In chapter 3, for instance, we run a logistic regression to assess the influence of different types of prior engagement on the probability of tie formation. Results are further corroborated using the Quadratic Assignment Procedure, which accounts for structural dependencies between relational data.

In chapter 4 we combine a standard econometric model with more recent techniques from the field of natural language processing. We use cosine similarity to construct a measure of thematic specialization based on the lexical overlap between project abstracts, and we run a beta regression to assess how partners' access to diverse sources of knowledge and ideas influences the observed degree of specialization in their repeated interactions.

1.5. Thesis structure

This dissertation consists of five chapters. The first chapter, the introduction, aims to contextualize the research conducted and motivate its contribution to the broader academic landscape. It contains a synopsis of the key theoretical frameworks underpinning the analysis as well as a list of 4 research questions, organized along 2 principal research lines. Additionally, to enhance overall clarity and readability, the introduction offers a concise graphical representation of the document's structure and flow.

Chapter 2 sets the stage by outlining the empirical setting upon which this dissertation is built. It delineates the data collection process and presents the key attributes of the Valencian policy-induced R&D network. The chapter begins by discussing both the rationale for policy intervention and the anticipated benefits of incentivizing collaborative R&D. It then describes the particular features that distinguish Valencia's adopted set of policy tools. More specifically, it shows that the instruments designed by the regional government gave rise to a network structure dominated by micro and small enterprises, with a well-defined core of highly active public research organizations. The observed characteristics of the network are important for understanding the empirical work undertaken in the next two chapters.

Chapter 3 looks at the structural evolution of the Valencian network, and the role of social proximity as a key driver of tie formation between regional actors. It explores empirically how prior experience in failed project applications shapes actors' collaborative behavior. The chapter compares the observed effect of "failure" to that generated by successful project applications to derive further insights. The moderating role of partners' cognitive proximity is also examined. The results validate the importance of social proximity in consolidating persistent ties. Yet, proximity generated by joint experience in failed applications was found to exert a stronger influence on tie formation than joint experience in successful applications. Moreover, the propensity of actors to re-engage following a rejection, was shown to be greater when they are cognitively distant. On a broader level, the analysis demonstrates, that in the absence of information on rejected projects, researchers in the field may end up overestimating the effect of prior successful experience.

Chapter 4 extends the analysis, presented in chapter 3, by delving into the implications of repeated collaborations. It therefore focuses exclusively on R&D partnerships which have received government funding and progressed to the project execution stage. The chapter investigates whether and under what condition organizations use recurring collaborations to explore new R&D topics (what we call diversification), as opposed to deepening their expertise in a single one (specialization). While chapter 3 confirms organizations' tendency to repeat collaboration with the same partner, chapter 4 sheds light on the practical implications of this behavior. It demonstrates empirically that strong ties are not always associated with the exploitation of the same topic, but that diversification in recurring collaborations is more likely when at least one of the partners involved mobilizes a diverse network of alters that could feed in novel knowledge and insights. The analysis underscores the significance of considering the function of strong bonds, given that organizational approach to repeated collaborations can be evidently distinct. Implications for knowledge transfer and overall performance are also explored.

The dissertation ends with a general discussion and conclusion. The implications of this doctoral work are elaborated with respect to the literature on R&D networks, and policymakers aiming to stimulate research interaction between regional innovation actors. Finally, limitations and research perspectives arising from this work are presented.

Figure 1.1 below offers a graphical representation of the structural organization and flow of the thesis.

CHAPTER 1

Introduction to the rise of collaborative R&D, academic interest in the structural dynamics of knowledge networks, contributions of this thesis to existing literature.

	CHAPTER 2	CHAPTER 3	CHAPTER 4
FRAMEWORK		*Input perspective*	*Output perspective*
		Types of social proximity and the origins of strong ties in knowledge networks	Implications of strong ties for knowledge transfer and thematic specialization
RESEARCH QUESTIONS	What are the main characteristics of the Valencian policy-induced R&D network?	1(a) How does joint experience in failed project applications influence future tie formation? 1(b) How does this relationship play out for partners with different levels of cognitive proximity?	2(a) To what extent do organizations leverage repeated collaborations to exploit the same topic multiple times (specialization) or to explore new ones (diversification)? 2(b) To what extent does partners' range of unique connections to other organizations inspire diversification in their repeated collaborations?
APPROACH	Descriptive-inductive	Hypothetical-deductive	Hypothetical-deductive
METHOD	Social network analysis	Logistic regression; Quadratic assignment procedure	Beta regression

CHAPTER 5

General conclusion and reflections on the main findings of the thesis, recommendations for policymakers, outline of promising trajectories for future research.

Figure 1.1. A graphical representation of the structure of the doctoral thesis.

Chapter 2. A closer look at the case of Valencia and its regional network of R&D partnerships

This chapter uses data from an online dashboard, developed as part of a three-month non-academic secondment with the Department of Innovation, Universities, Science and Digital Societies within the regional Valencian Government. Available at: https://yankovadn.github.io/RIS3CV/

2.1. Introduction

The goal of this chapter is to introduce the inter-organizational R&D network in the Valencian region, which emerged in response to targeted policy incentives, integrated in the regional smart specialization strategy. This network provides an overall picture of the distinctive attributes of the Valencian innovation system and constitutes the core empirical setting upon which this dissertation is built. For these reasons, it is essential to present a detailed account of this R&D network. The chapter is primarily descriptive in nature and ought to be seen as a prelude to the quantitative analysis, contained in the following two chapters.

To begin, the chapter provides a concise contextualization of the growing emphasis on networking and R&D partnerships, particularly in the context of regional innovation systems (section 2.2). It highlights both the rationale for policy intervention and the anticipated benefits of incentivizing collaborative R&D. Smart specialization strategies and the Partnerships for Regional Innovation are only some of the most recent examples of such type of policies. Next, the chapter presents the case of Valencia and its institutional setting. A moderate innovator, according to the 2023 European Innovation scoreboard, Valencia is characterized by a highly fragmented business fabric and a strong predominance of micro and small enterprises, which are only weakly connected to the regional network of universities and research centers. All of this renders Valencia comparable to many other regions in southern and eastern Europe, adding greater relevance to the findings, presented later on in the thesis. Section 2.3 also provides a brief overview of Valencia's smart specialization strategy and the specific policy instruments developed by the local authorities to strengthen intra-regional partnerships. Section 2.4 delineates the data-collection process, which gathered detailed information on a total of 444 R&D partnerships over a 7-year time period: from 2016 to 2022. This dataset serves as the basis for the empirical analysis carried out in the next two chapters of the thesis. We also present a concise spatial and a-spatial analysis of the resulting R&D network, applying several techniques from social network analysis. The key features of the network are highlighted alongside the most important players in the regional context and the main patterns of collaboration. The last section summarizes the principal takeaways and provides a segue to the next chapter.

2.2. Policy-induced R&D networks: rationale for intervention and expected benefits

The past two decades have witnessed a growing support for knowledge networks and partnership-based R&D (Doern & Stoney, 2009; Martin, 2016). The policy mix across European countries has diversified to provide extra incentives for interaction between heterogeneous organizations, including firms, universities, research centers and others (Cunningham et al., 2013; Cunningham & Gök, 2012). The original focus on national innovation systems (Lundvall, 1992; Nelson, 2004; OECD, 1997) was gradually complemented by a set of regional and sectoral policies, inspired by concepts like learning regions, innovative milieu, industrial districts and clusters (Camagni, 1995; Cooke, 2003; Florida, 1995; Morgan, 1997; Porter, 2000; Simmie, 2011). In this regional innovation system (RIS) logic, cooperation between firms and knowledge-creating and diffusing organizations remains ever more present as a target of policy intervention (Cooke, 2008; Doloreux & Parto, 2005).

Some of the more recent examples of RIS policies include the smart specialization (S3) paradigm (Foray et al., 2009, 2018), which is arguably the most central pillar of European Cohesion Policy, and the Partnerships for Regional Innovation (PRI) (Pontikakis et al., 2022), which build on the lessons learned during the first seven years of S3 implementation. At its core, smart specialization is about uncovering locally embedded assets, capabilities and competences and prioritizing a few strategic domains, where regional competitiveness can be achieved (Foray, 2017). The PRI concept, on the other hand, puts greater focus on directionality. It aims to leverage the innovative potential of regions toward accelerating social and sustainable transformations (Pontikakis et al., 2022). Ultimately, both approaches recognize that system-level changes require the involvement of many different actors and the construction of collaborative links and partnerships between them.

From a policy perspective, the rationale for government intervention in stimulating such partnerships is twofold. On a conceptual level, the wicked nature of today's societal challenges clearly requires joint efforts by previously isolated silos of innovating actors. As pointed out by Cunningham & Ramlogan (2012), the market is no longer looking for a single technology or one-off service, but for systemic innovations in the form of integrated "packaged solutions". They further point out that such packaged solutions can only be developed and produced by networks of closely-connected actors with diverse knowledge and competences. On a more practical level, government intervention is needed to fix persistent failures associated with limited access to information and insufficient awareness of the benefits of collaboration, little willingness to engage in knowledge transfer activities with other actors, as well as overall weak internal structures.

If successful, government interventions in the form of targeted policy instruments, can generate a number of positive externalities, which may not have occurred otherwise. Following the additionality framework[2] (Georghiou, 2002), firms may be more likely to engage in cooperative agreements and better prepared to appropriate the benefits from them, if relevant public incentives are in place (Luukkonen, 2000). Drawing on existing literature, the anticipated benefits of government-supported R&D networks can be summarized as follows: (i) attaining economies of scale and scope, (ii) enhanced competitiveness due to better connectivity and foresight of stakeholders, (iii) faster and more efficient technology transfer, (iv) new capability creation as a result of inter-organizational learning and the exploitation of complementary resources, (v) decreasing R&D costs due to pooling risks and co-opting competition (Hagedoorn et al., 2000; Martínez-Noya & Narula, 2018).

In sum, the increasing focus on regional systems of innovation, coupled with the introduction of policies explicitly prioritizing network formation and inter-organizational partnerships, requires greater scrutiny of the structure and dynamics of the emerging

[2] Within the broad framework of behavioral additionality, some specific concepts have also emerged. For instance, "network additionality" is understood as the ability of public funding instruments to increase interaction and cooperation between organizations to a greater extent than would be present without such funding (Hyvärinen & Rautiainen, 2007; Rossi et al., 2016).

collaborative networks. The subsequent section contextualizes the case of Valencia and its regional R&D ecosystem, which serves as the primary empirical setting for the analysis presented in this thesis.

2.3. The case of Valencia and its smart specialization strategy

Valencia is among the more advanced regions in Spain and the fifth most innovative, according to the 2023 EU Regional Innovation Scoreboard[3]. In 2014 the regional government approved the Research and Innovation Smart Specialisation Strategy (RIS3-CV) of Valencia, although implementation of the document only began around 2016. As pointed out earlier, the smart specialization paradigm represents a marked shift in EU innovation policy, as it encourages regions to undertake a bottom-up exploration of their local assets – technological and scientific domains – and translate those into future competitive advantage (Foray et al., 2009). Following this framework, the RIS3-CV strategy defined several priority axes, which essentially provide a roadmap for the design of public R&D policies and actions in the region. The total budget allocated in the strategy for the 2014-2023 period exceeds 1.6 billion EUR[4], making it one of the most important public policy instruments for stimulating regional research and innovation. This is why in this dissertation we rely on the implementation of the RIS3-CV strategy to obtain a reasonable picture of the local publicly-funded R&D system and the underlying knowledge network.

It is important to note that Valencia is characterized by a highly fragmented business fabric. More than 99% of the firms registered in the region are classified as either micro, small or medium enterprises. Due to their size, many of them find it difficult to compete for external funding or support that would allow them to develop new products and processes (Generalitat-Valenciana, 2016). This generates a vicious cycle whereby firms which are arguably most in need of external support for innovation are also the ones least capable of absorbing public funds earmarked for them – a situation commonly known as "the regional innovation paradox" (Oughton et al., 2002). In addition, the limited culture of collaboration, especially among micro and small enterprises, further hinders their potential to increase productivity. Recognizing this as a key challenge, Valencia's smart specialization strategy aims to not only steer innovation in prioritized areas, but also to stimulate interaction between weakly-connected nodes, especially between firms, knowledge-producing and knowledge-diffusing organizations.

In Valencia the management of the individual policy instruments, laid down in the strategy, was distributed between several local agencies. Two of them, however, played a leading role: the Valencian Institute for Business Competitiveness[5] (IVACE) and the Valencian Innovation Agency[6] (AVI) (Generalitat-Valenciana, 2019). IVACE was established in 1984 and its mission is geared toward assisting regional SMEs in increasing their competitiveness and

[3] The European innovation scoreboard interactive tool is available at: https://ec.europa.eu/research-and-innovation/en/statistics/performance-indicators/european-innovation-scoreboard/eis

[4] The total budget of 1.6 billion Euros combines funding from regional, national and European sources.

[5] The official website of IVACE can be accessed through the following URL: https://www.ivace.es/index.php/es/

[6] The official website of AVI can be accessed through the following URL: https://innoavi.es/en/

overall innovative capacity. AVI, on the other hand, was created more recently in 2018, specifically for the purpose of managing the innovation strategy of Valencia and improving the regional productive model. Together, these two organizations administer approximately 75% of the 1.6 billion EUR that the regional government has designated for the implementation of the local innovation strategy (Generalitat-Valenciana, 2019).

With regards to promoting collaborative R&D, AVI has two lines of support: (1) **Strategic projects in cooperation** and (2) **Consolidation of the business value chain.** The second program accepts individual projects as well, but given the focus of this research, we have concentrated exclusively on collaborative ones. IVACE has one consistent program dedicated to collaborative R&D, called (3) **R&D in cooperation**. The table below provides detailed information on the characteristics of each policy instrument.

Table 2.1. Description and overview of the three programs.

Program characteristics	R&D in cooperation	Strategic projects in cooperation	Consolidation of the business value chain
Funding agency	IVACE	AVI	AVI
Project typology	downstream R&D projects in industrial research; experimental development	downstream R&D projects in industrial research; experimental development; process, organizational and product innovation	downstream R&D projects in industrial research; experimental development; process, organizational and product innovation
Eligible entities	Private firms	Private firms, universities, research centers, hospitals, non-profits involved in R&D activity	Private firms
Mandatory partners	At least 1 SME	At least 1 public research organization (PRO)[7], though it does not have to appear as a partner in the consortium (sub-contracting is allowed)	N/A
Entities' location	Region of Valencia	Region of Valencia	Region of Valencia
Funding size	80 000-500 000 EUR	>500 000 EUR	>500 000 EUR
Min co-financing base	15% (for all partners)	15% (for all partners)	15% (for all partners)
Project duration	1-2 years	2-3 years	2-3 years
Evaluation criteria	(i) proposal quality, (ii) technical and financial capacity of the team members to carry out the project, (iii) characteristics of the firms, (iv) anticipated project impact and alignment with the regional smart specialisation strategy, social and environmental goals	(i) proposal quality, (ii) technical and financial capacity of the team members to carry out the project, (iii) complementarity of the team, (iv) anticipated project impact and alignment with the regional smart specialisation strategy, social and environmental goals	(i) proposal quality, (ii) technical and financial capacity of the team members to carry out the project, (iii) complementarity of the team, (iv) alignment with social and environmental goals

[7] In this document the term PRO encompasses universities, research centers and technological institutes.

All three programs were financed by a mixture of local funds and ERDF funds, and all three run on an annual basis. This means that every year, AVI and IVACE allocate the available funding on a competitive basis to a selected number of consortia – purposefully created and legally binding groups of collaborating partners.

When looking at the three policy instruments together, the following aspects stand out:

(i) There is a clear emphasis on short-term applied research, rather than basic or more explorative research. This is consistent with prior observations that national or EU-level programs in general tend to support pre-competitive research, while regional instruments are more likely to prioritize research that is closer to market (Hagedoorn et al., 2000).

(ii) There are few signs of sectoral preferences or prioritization, despite the soft incentives in the evaluation criteria regarding "alignment with the regional smart specialization strategy". In that sense, the local government is refraining from direct intervention and is not placing any significant restrictions on the potential topics of R&D cooperation.

(iii) There is a pronounced regional focus, as all partners are expected to be located in Valencia. Extra-regional partnerships do not fall within the scope of either of the three policy instruments. We can, therefore, conclude that the implicit objective of the programs is the strengthening of the intra-regional network of R&D partnerships, rather than the explicit "import" of external competencies, technologies or skills.

(iv) There is no "minimum consortium size" requirement in any of the instruments. This suggests that policymakers were more interested in stimulating firm-firm and firm-PRO relationships, than in broadening the connections companies have (i.e. expanding the size of their network of contacts).

In sum, despite some differences in eligibility criteria, all three programs aim to enhance cooperation between local stakeholders and to support innovative activities related to the creation of new products, processes or services in the Valencian region.

Figure 2.1 shows the distribution of project applications per year. It is visible that during 2016 and 2017, public funding for collaborative R&D projects under the RIS3 strategy was mostly managed by IVACE, but the creation of AVI in 2018 opened up new lines of financial support for cooperation between regional actors. Approval rates do not fluctuate significantly between years.

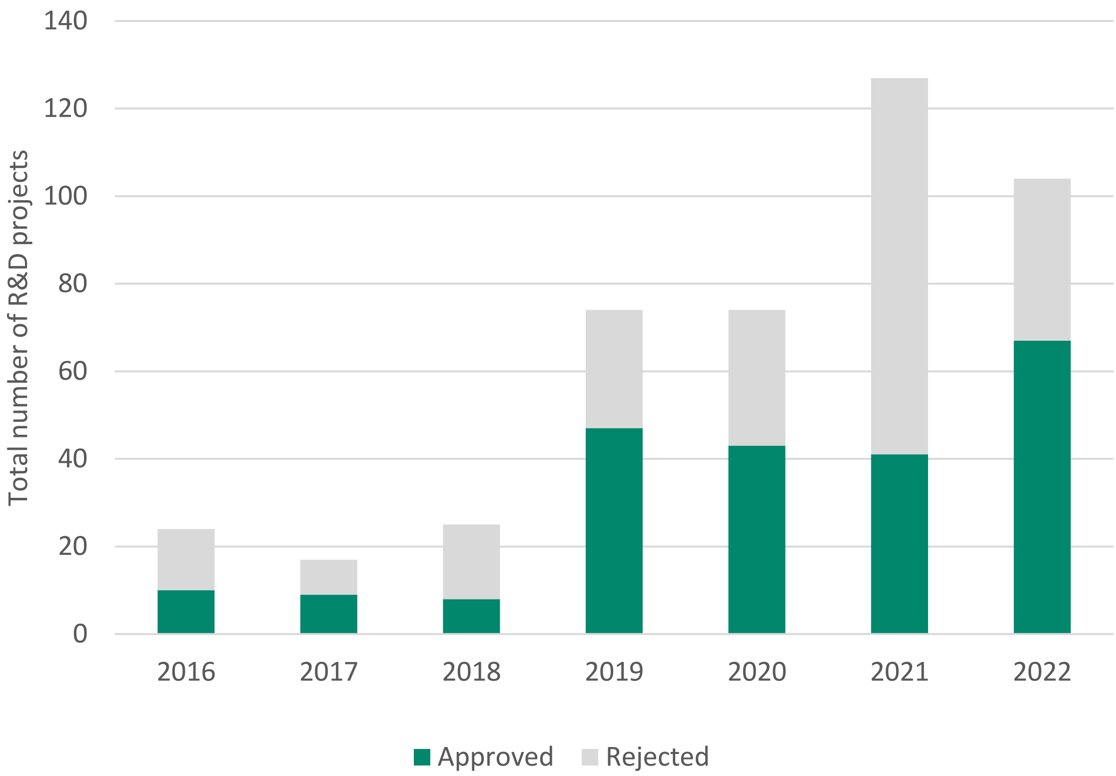

Figure 2.1. Distribution of approved and rejected R&D project applications per year.

2.4. Data collection and overview

The data collection process relied on public records, published on the websites of AVI and IVACE. AVI provides information on both approved and rejected R&D partnerships, while IVACE only discloses the approved ones. Through consultations with public officials, however, a complete list of rejected projects was also obtained. To be clear, the term "rejected" here refers to R&D partnerships which were not awarded government funding for the execution of their project proposal. The "approved" ones, on the other hand, successfully passed the evaluation stage and entered into project execution.

The final longitudinal dataset includes all R&D partnerships (approved *and* rejected), which applied for public funding through one of the three policy programs between 2016 (when the implementation of the regional smart specialization strategy *de facto* began) and 2022 (when the most recent information was uploaded online). Public records include the name of each project, as well as the name and CIF[8] of all partners. For approved projects, the individual financial subsidy allocated to the partners is also made available. Thus, we can contend that the information captures the full range of actors who responded to government-placed incentives and formed an R&D partnership, irrespective of these partnerships' success in obtaining public subsidy (i.e. regardless of the funding preferences of the two agencies AVI and IVACE). This unique feature of the dataset is exploited empirically in chapter 3 of this thesis.

[8] The CIF number is a unique combination of letters and numbers used to identify a Spanish company or legal entity for tax purposes.

In the 7-year time window, the total number of R&D partnerships amounts to 444. Roughly half of them received public funding, while the other half were rejected. Project team size varies between 2 and 9 partners, with an average of 3 organizations per team. In this document universities, research institutes, both general and health-related, as well as technological institutes are all considered public research organizations (PRO). The table below provides additional descriptive information.

Table 2.2. Descriptive information on projects and organizations involved.

Total number of projects	**444**
Share of approved projects	50.7 %
Min number of partners per project	2
Average number of partners per project	3
Max number of partners per project	9
Total number of organizations	**554**
Private for-profits[9]	498
Micro Enterprise	*178*
Small enterprise	*151*
Medium-sized enterprise	*98*
Large enterprise	*71*
Associations	12
Technological institutes	11
Cooperatives	9
Universities	7
Research institutes	7
Health research institutes	7
Others	3

During the period of observation, a total of 554 regional organizations applied to one of the three programs as part of a team. As in many collaborative networks, 90% of those applicants were private for-profits. The SABI database[10] was used to match companies by their unique identifier, and retrieve information on their age, size, geographic location and economic activity. We analyzed the characteristics of these companies in more detail in order to better understand their profile.

36% of program participants are microenterprises with a turnover below 2 million EUR, and 67% have 50 employees or less, which is highly illustrative of the Valencian context. It also implies that smaller regional players, which generally lack sufficient resources to compete in excellence programs (i.e. Horizon 2020), are in fact quite active in the local policy-induced R&D network. The average age of all 498 enterprises is 24 years, while the median – 19. The youngest firm was only 1 year in business when it applied for funding, while others, founded at the end of the XIX century, were more than 100 years old. Hence, the dataset is characterized by a mix of young start-up like enterprises and older more established firms.

[9] Based on EU recommendation 2003/361: - Micro enterprises: ≤ € 2 m turnover, - Small enterprises: € 2 -10 m turnover, - Medium-sized enterprises: € 10 - 50 m turnover, and - Large enterprises: ≥ € 50 m turnover.

[10] SABI (Iberian Balance Sheet Analysis System) is a database with financial information of more than 2,6 million companies in Spain and Portugal. 100% of the private companies in our sample were successfully matched with this database.

We also examined their economic activity based on the NACE classification. “C. Manufacturing”, “J. Professional, scientific and technical activities” and “M. Information and Communication” accounted for 80% of companies’ registered economic activity. Across the Valencian region, however, these three industry categories are less strongly represented, as they account for roughly 28% of local employment and 19% of all registered businesses[11]. In simpler terms, while the companies applying for collaborative R&D funding hailed from industry sectors with high R&D intensity, they do not fully mirror the broader industry composition of the region. For each of the three NACE categories (C., J. and M.), we singled out the most common classes of activity, which were relevant to at least 10 companies. Table 2.3 shows the number of firms per industry class. A complete account of all companies’ registered type of economic activity is available in the Appendix (2.A).

Table 2.3. Industry categories with greater representation in the regional Valencian R&D network.

Industry category	Description of the two-digit NACE Rev.2	Num. of firms
C. Manufacturing	22. Rubber and plastic products	36
	13. Textiles	32
	20. Chemicals and chemical products	28
	28. Machinery and equipment	20
	25. Fabricated metal products, except machinery & equipment	16
	23. Other non-metallic mineral products	14
J. Information & Communication	62. Computer programming, consultancy and related activities	49
M. Professional, scientific & technical activities	71. Architectural and engineering activities	52
	72. Scientific research and development	31

Note: Scientific research and development refers exclusively to biotechnology, natural sciences and engineering.

Despite strong empirical evidence that R&D collaboration is far more common in high-tech sectors (Hagedoorn, 2002), when looking at the publicly subsidized partnerships in a moderate innovator region like Valencia, we observe a more balanced involvement of both medium-high tech companies[12] (producing chemical products, machinery and equipment), and medium-low to low tech companies specialized in the manufacturing of rubber, plastic products and textiles. It is also worth noting that the last three subcategories belong to what is formally known as KIBS (Knowledge Intensive Business Services). The increasing involvement of KIBS firms in regional innovation systems as a type of boundary spanners between the environment and SMEs has been well-documented in the literature (Muller & Zenker, 2001). KIBS seem to assume an interface function by reinforcing and catalyzing the innovation capacities of their partner SMEs, especially if they are underperforming or lagging behind. A similar dynamic seems to be at play in the case of Valencia’s regional system of R&D partnerships.

Aside from the private for-profit organizations, the table of participating entities (table 2.2) also features 11 technological institutes. Those can be considered a unique element of the local innovation ecosystem, and as such, they also require some contextualization. Established with support from regional business associations and the government between

[11] Employment data and data on number of companies per industry in the Valencian region were collected from the Spanish National Statistical Institute for the year 2020.
[12] Based on Eurostats’ classification of the technological intensity of manufacturing industries.

the 1970s and the 1990s, the institutes operate as private research non-profit entities, whose primary goal is to support regional SMEs in advancing their capacities and innovative activity. Each institute is housed in a single geographic location (i.e. no dispersion of research activity), and is generally dedicated to a specific topic, such as energy, textile, biomechanics, or others[13].

Valencia also has several independent research centers, which do not belong to a university structure. Some of them work exclusively on health-related topics. These are the so-called "Health research institutes". Finally, we also note in table 2.2 the presence of Cooperatives and Associations, whose role in the innovation process is often overlooked.

2.5. Analysis of the regional network of R&D partnerships in Valencia

This section presents an overview of the resulting inter-organizational network, mapped both spatially and a-spatially in the 7-year time window (2016-2022). The use of social network analysis, a method rooted in graph theory, allows us to better characterize the patterns of relationship between the relevant actors, and the implications of those links for the structure and dynamics of the network.

The application of SNA in economic geography, and especially in the study of regional innovation processes, has surged in the past decades (Glückler & Doreian, 2016; Ter Wal & Boschma, 2009a). In SNA, each unique organization is represented as a *node* or *a vertex*. Two nodes are connected by a line (*edge*) when they have some type of formal or informal relationship. In this case, two organizations are connected when they are members of the same R&D partnership. Thus, we are in practice projecting a two-mode network of type "project – organization" into a one-mode network, whereby each node is a separate organization and each link implies membership in the same R&D partnership (Borgatti, 2009). In weighted networks, the frequency of interaction is also reflected, either through the thickness of the connecting line, or the location of nodes in space (more frequent partners are mapped closer to each other). In our case, the frequency of interaction is based on the number of projects the two partners have jointly participated in.

Figure 2.2 illustrates the policy-induced network of relationships between all 554 Valencian organizations in the time period 2016-2022. The network is presented a-spatially, which facilitates the identification of topological features and the relative position of individual organizations. A closer look at the overall degree distribution shows a power-law pattern, although not a perfectly conforming one (α = 1.6). The suggests the presence of a dense core of tightly connected actors, also known as hubs (Barabási, 2009; Barabási & Bonabeau, 2003).

To better assess the influence of each entity and the nature of these hubs, a degree centrality measure was estimated for all nodes in the network. Degree centrality is simply a count of how many unique partners a single node (the *ego*) has. It can also be understood as the size of the ego´s network. Thus, for instance, an organization which has participated

[13] Information on the 11 institutes can be found on the REDIT website: www.redit.es

in 2 distinct projects with a total of 7 different partners will have a degree centrality of 7. Degree centrality is often used as a proxy for the importance of a given node (Freeman, 1978). It signals a dual advantage: (i) the number of partners the ego can source novel knowledge and resources from, and (ii) the number of partners the ego can influence directly without passing through intermediaries.

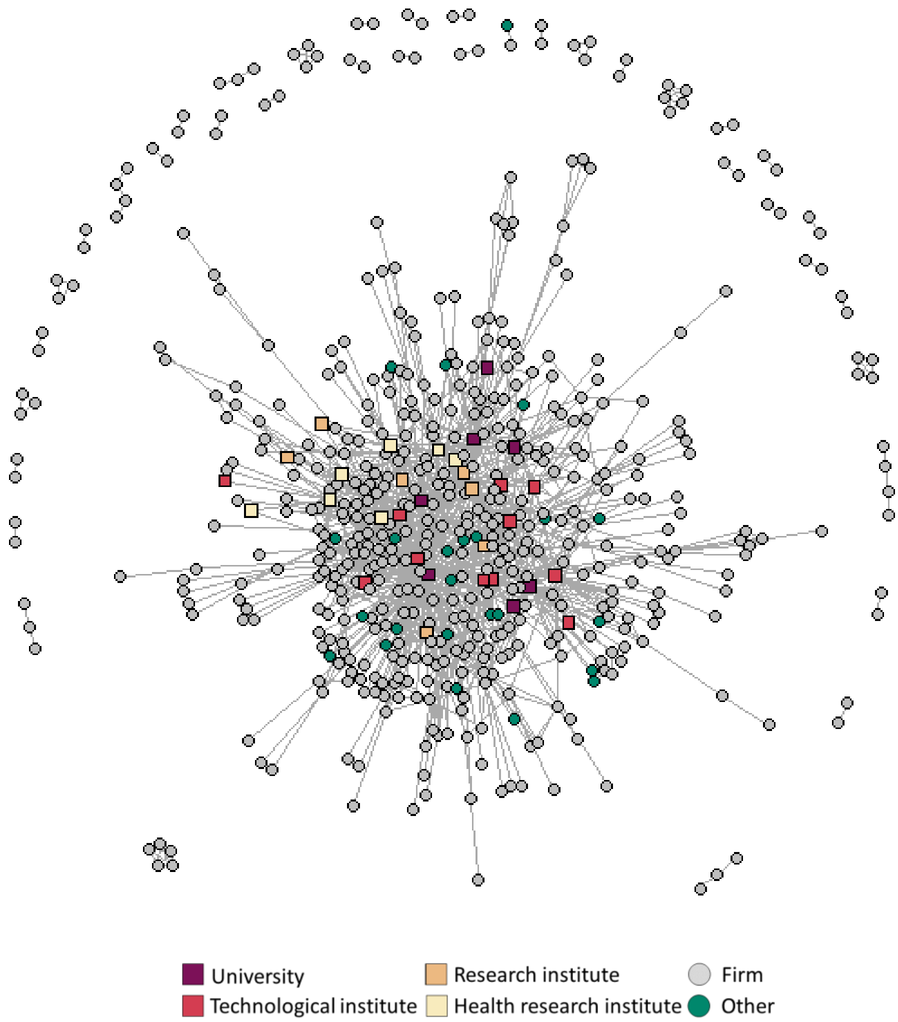

Figure 2.2. An aspatial map of the regional R&D partnerships in Valencia between 2016 and 2022 (funded and unfunded combined). *The network was generated using R, with a Fruchterman-Reingold layout algorithm. This is a force-directed layout that simulates the presence of an attractive force between connected vertex pairs. The more often two nodes collaborate, the closer they will appear to each other on the graph.*

Figure 2.3 shows the distribution of degree centrality for all 554 nodes in the network. As could be expected in scale-free networks, the distribution is highly skewed, with only a handful of organizations serving as hubs that assist the connectivity of the entire network (Barabási & Bonabeau, 2003). The plot confirms that the actors with highest degree centrality in the network are PROs. This is consistent with other studies of regional and extra-regional networks (Expósito-Langa & Molina-Morales, 2010; Roediger-Schluga & Barber, 2006), where universities and large research centers were found to serve as intermediaries, and thus appear as frequent participants in regional partnerships. Similarly, in the case of Valencia, during the 7 years of observations, the Polytechnic University of Valencia established R&D partnerships with 178 different organizations. Among those are firms, other universities, research center and technological institutes. Although not all of those partnerships were funded, the joint application implies the formation of some type of inter-organizational relationship.

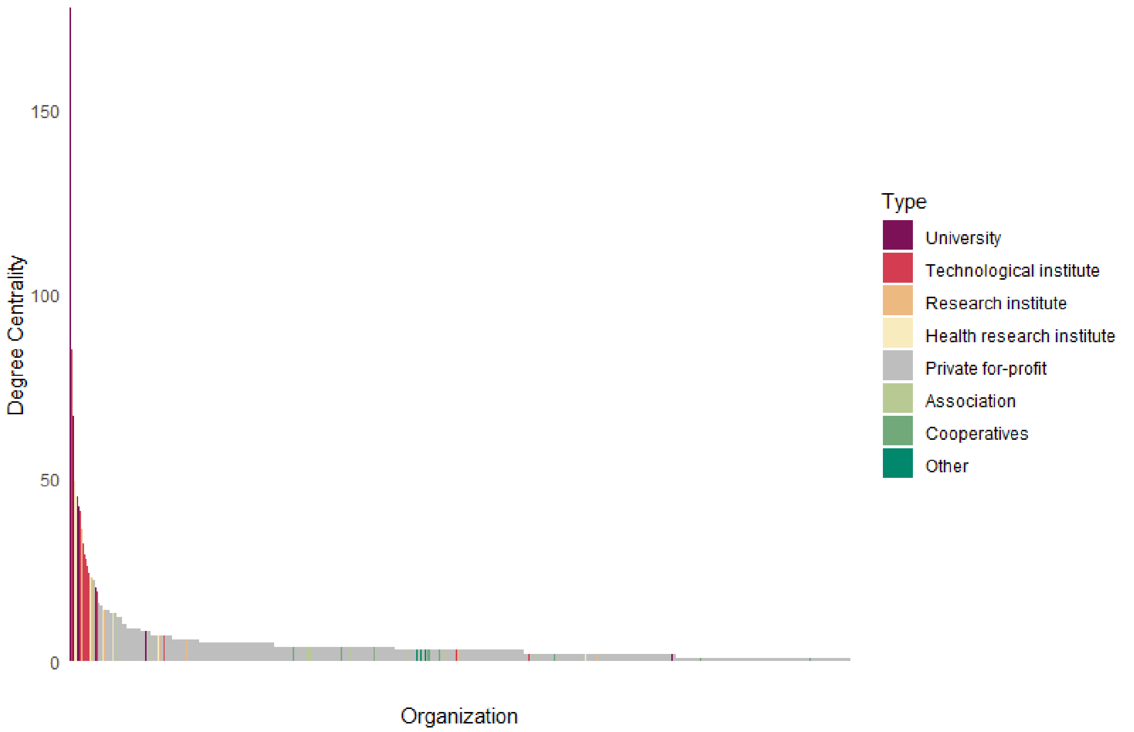

Figure 2.3. Distribution of degree centrality per organization in the Valencian inter-organizational network.

In prior studies, the prominent role played by PROs has been attributed to their outward knowledge-disseminating orientation. Larger universities and research centers normally possess the human, technical and financial resources to maintain links with many organizations at the same time. Private companies, on the other hand, especially R&D intensive ones, are more concerned with privacy and data protection and are generally more cautious to avoid potential opportunistic behavior from partner firms. In addition, many of those classified as micro or small enterprises may simply lack the resources to maintain more than one partner relationship. Therefore, the majority of companies can be expected to realize only a few strategic partnerships.

In a next step, we map the same network of collaborative relationships, except this time spatially, across the geographical area of Valencia region. The location of each organization was geo-coded using the longitude and latitude coordinates of its registered address. Figure 2.4 (a) depicts the spatial distribution of the main PROs in the region, namely 5 universities and 11 technological institutes. Some of them are part of the same campus or science park, and their geographical coordinates may therefore overlap. The map is also color-coded to highlight the three main cities: Valencia with a metropolitan area encompassing some 800 000 inhabitants, Alicante – a medium-sized city with 330 000 inhabitants, followed by Castellon with 170 000. Figure 2.4 (b) contains the network of relationships between all actors, giving rise to several noteworthy observations.

First, while participation in R&D partnerships is recorded from actors across the region, there is a significant spatial concentration in the vicinity of the three main economic centers. They appear to serve as dynamic innovation hubs that actively interact with each other. Second, while Valencia city boasts a dense internal network of inter-organizational relations, the dynamics in Castellon and Alicante are curiously distinct. The latter two display a more pronounced pattern of interactions between each other, but do not exhibit similar levels of connectivity *within* their own urban boundaries. In other words, there seems to be more inter- rather than intra-city connectivity. Thus, we can say that the policy instruments, implemented by the local government, manage to consolidate tight links between the three primary economic centers of the region, but the impact is barely extended to the peripheral municipalities. Most of those municipalities do not form part of the network at all. The overall density of the entire network is only 0.009, which means that roughly 0.9% of all possible partnerships between the 554 participating organizations were realized.

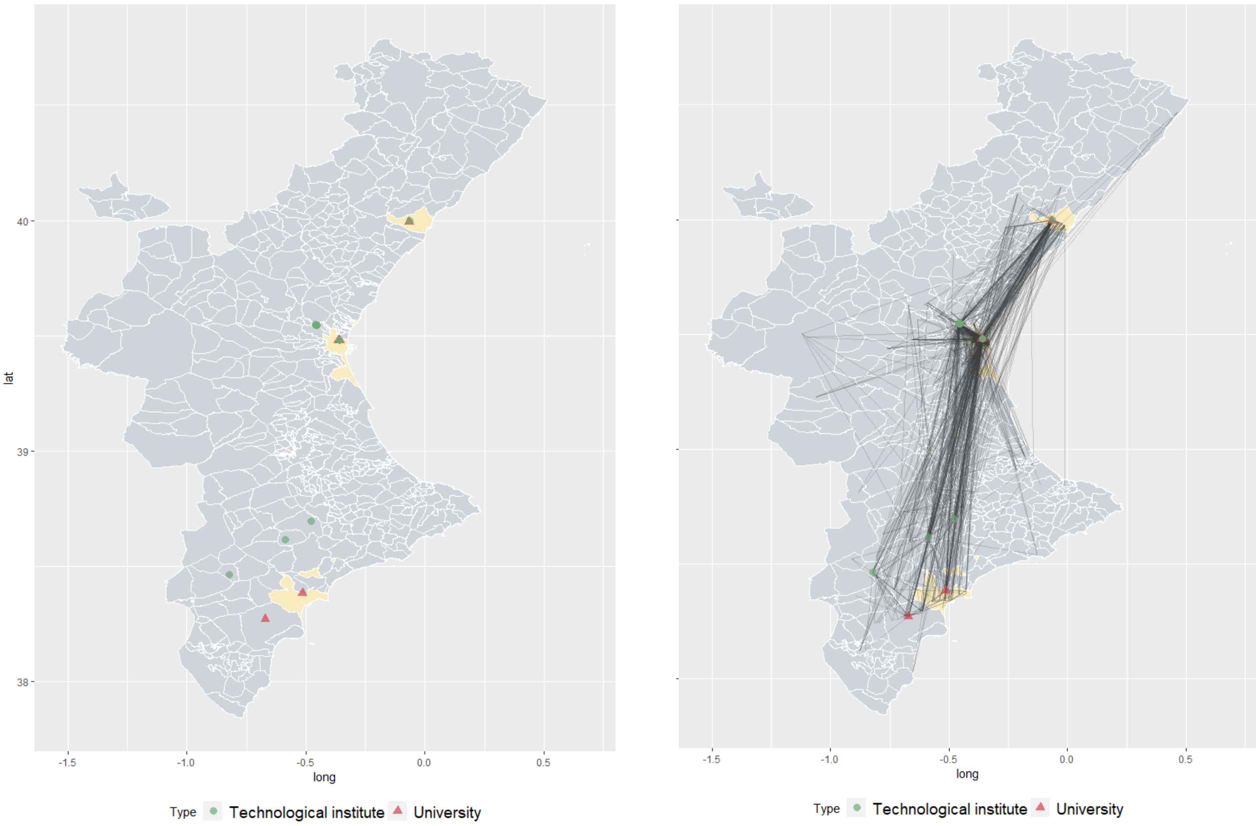

Figure 2.4 (a). A geographical map of the Spanish region of Valencia. The main universities and technological institutes are marked with a red triangle and a green circle respectively. Colored municipalities from top to bottom: Castellon, Valencia, Alicante. (b) A spatial mapping of the policy-induced knowledge network between regional actors in the time period 2016-2022.

To further explore the spatial distribution of partners and how it relates to their collaborative choices, we plot the straight-line geographical distance between all pairs of partners in the network. Figure 2.5 shows the resulting graph. The maximum spatial separation between two partners in the network is 205km, while the shortest is practically 0km and it denotes instances when the partners are located next-door to each other or form part of the same campus or cluster. It is also worth noting that more than half of all organization pairs are located less than 25 km from each other, and the median value is only 14 km. While the observed spatial proximity between some partners may signal that they are members of the same cluster, it could also imply that the organizations were familiar with each other from before. In this context, further research is needed to determine the extent to which the policy instruments implemented by AVI and IVACE actually stimulated the formation of new links as opposed to reinforcing existing network connections. Figure 2.5 also suggests a pronounced preference among participating entities to collaborate with near-by actors – a common phenomenon in knowledge networks (Broekel & Boschma, 2012; Tsouri, 2019; Vicente et al., 2011).

Overall, we can conclude that the spatial concentration of R&D partnerships near the main economic centers highlights their role as innovation hubs in Valencia, while the preference for proximate partners underscores the significance of existing connections.

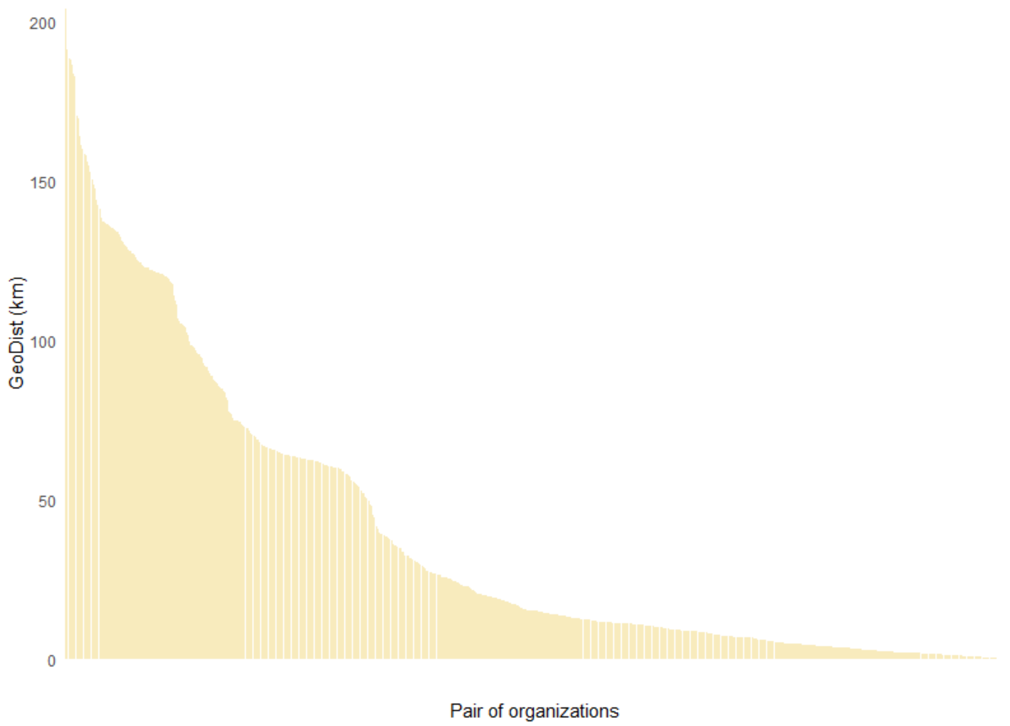

Figure 2.5 Distribution of the geographical distance between each unique pair of partners in the network.

2.6. Conclusion and key takeaways

The goal of the present chapter was to introduce the empirical setting of this dissertation, namely the case of Valencia's regional network of publicly-funded R&D partnerships in the time period 2016-2022. It contextualized the policies implemented by the local government to incentivize inter-organizational interaction. The analysis of AVI and IVACE´s three policy programs produced several important insights regarding the structure and dynamics of Valencia's regional R&D network: (1) All three policy instruments were geared toward strengthening intra-regional synergies and cooperation, with a clear focus on downstream applied research. (2) The emerging R&D network was dominated by micro and small enterprises, which fits the overall profile of Valencia as a moderately innovating region with a fragmented business fabric. It also shows that the targeted policy programs managed to reach smaller regional actors, who are often unable to compete in large-scale excellence programs on a national or international level, such as the EU-FP ones. (3) The network of partnerships is sparse and weakly scale-free. It is characterized by a few well-connected hubs, comprised mostly of universities and research centers. Most notable is the role played by the Polytechnic University of Valencia, and the local technological institutes. The network is spatially concentrated along the three main cities: Valencia, Alicante and Castellon, and barely extends to smaller municipalities inland. These observations are important to understand and contextualize the analysis presented in the following two chapters.

Chapter 3. The effect of failed R&D project applications on organizations' collaborative behavior

This chapter is based on the paper:

Yankova, D., D'Este, P. & García-Melón, M. (2023) "Towards a New Facet of Social Proximity: The Effect of Failed Project Applications on Organizations' Collaborative Behavior in R&D Networks". Working paper version available at SSRN: https://ssrn.com/abstract=4478790 or http://dx.doi.org/10.2139/ssrn.4478790
[Submitted to "Research Policy" journal; status: under review]

Abstract

Prior engagement between partners is a well-established antecedent of tie formation in interorganizational networks, and it is frequently operationalized in the context of successful collaborations that have received financial support. In reality, in competitive publicly-funded R&D networks, many partnerships never materialize into actual collaborations and are rejected in the application stage. Yet, we know very little about how failed applications shape organizations' collaborative behavior and the R&D network as a whole. In this chapter we investigate to what extent joint experience in rejected project applications influences tie formation, and how this relationship plays out for partners with different levels of cognitive proximity. The empirical analysis is based on a policy-induced collaborative network of both approved and rejected R&D projects in the Spanish region of Valencia in the 2016-2022 time period. Results indicate that joint experience in failed applications exerts a positive influence on future tie formation, and a stronger one than joint experience in successful applications. Moreover, the propensity of actors to re-engage following a rejection, is greater when they are cognitively distant. On a broader level, the study demonstrates, rather critically, that in the absence of information on rejected projects, empirical studies may end up overestimating the effect of prior successful collaborative experience.

3.1. Introduction

Organizations with a history of prior engagement seem to gravitate toward each other on the basis of previously established mutual trust, shared norms and expectations (Gulati, 1995; Gulati & Gargiulo, 1999; Sorenson et al., 2006; Zaheer et al., 1998). However, defining what exactly constitutes prior engagement is a contentious issue. Are partnership failures similar to successful ones in terms of building social proximity and stimulating tie formation? Our study addresses this question in the context of publicly funded programs to support collaborative R&D networks by looking at the role of two types of prior engagement: partnership in failed vs. successful research project applications. Here "failed" refers to partnerships which were not awarded government funding for the execution of their project proposal.

So far, the vast majority of studies assessing the evolution of R&D networks only consider successful collaborations, that have received financial support (Autant-Bernard et al., 2007; Caloffi et al., 2015; D'Este et al., 2013; Heringa et al., 2016; Paier & Scherngell, 2011). In reality, in competitive publicly-funded R&D programs many partnerships never materialize into actual collaborations and are rejected in the application stage for a variety of reasons, ranging from project quality to misalignment with policy priorities or evaluation criteria. For comparison, Horizon 2020 – the EU's flagship R&I funding program – has an average success rate of 24% even after excluding ineligible and low-quality submissions (McCarthy, 2017). Despite the overwhelming prevalence of unsuccessful partnerships, this form of engagement between organizations is typically undetected and therefore – underexplored. We know very little about how failed applications shape organizations' collaborative behavior and the R&D network as a whole.

On the one hand, transaction cost theory suggests that organizations with joint experience in failed project applications may be particularly motivated to re-engage as a way of recovering some of the original hidden costs associated with searching for partners, negotiating and consolidating a working relationship (Amoroso, 2014; Takalo et al., 2013). At the same time, rejection can push organizations to pursue substitute partners, especially if they attribute the negative outcome to consortia composition. Hence, it is unclear whether failed project applications end up reinforcing existing patterns of collaborative behavior, or on the contrary: serve to reverse them. In that sense, relying on successful partnerships alone can provide only partial insights into how publicly funded R&D network structures come into being. In the absence of information on rejected project applications, researchers may end up attributing some portion of newly formed ties to chance, when they could in fact be driven by undetected instances of unsuccessful prior partnerships. This would suggest that we are underestimating the effect of relational embeddedness or social proximity. Alternatively, if failed partnerships decrease the probability of future tie formation, then we may conclude that relational embeddedness does not have a uniform effect and the type of past experience matters.

To our knowledge, so far only a handful of empirical studies have considered the combination of approved and rejected projects, and all of them test for the effect of organizations' characteristics – size, reputation, network position – or overall consortia composition on the probability of grant award (Barajas & Huergo, 2010; Enger, 2018; Enger & Castellacci, 2016; Wanzenböck et al., 2020). They do not explicitly explore how the outcome of the evaluation process (success vs. failure) influences network dynamics and subsequent partner selection. In this chapter we seek to address this gap by investigating (1) how joint experience in rejected project applications – a critical but overlooked element of social proximity – influences tie formation, and (2) how this relationship plays out for partners with different levels of cognitive proximity.

Hence, our work makes three contributions. Conceptually, it introduces a new facet of relational embeddedness, based on joint failure, thus challenging existing assumptions on what type of prior interaction is necessary to stimulate further engagement. Second, from an empirical standpoint, the study introduces an important omitted variable – past experience in failed project applications, which is arguably far more prevalent than joint experience in successful project applications. This allows us to get a more accurate picture and more robust estimates of the factors shaping tie formation. Finally, from a policy perspective, the chapter offers original evidence on the unobserved effects of public instruments targeting R&D collaboration, including the emergence of informal links between unsuccessful partners. Policymakers may thus be interested in how actors decide to readjust their partner selection process in response to failure to obtain competitive public funding for a joint project proposal.

The chapter is structured as follows: we first review the literature on social proximity in collaborative R&D networks, building up two opposing scenarios regarding the anticipated relationship between project rejection and future tie formation (section 3.2). Next, an overview of the methodological approach is provided. Results are presented in section 3.4, while the conclusion (3.5) offers a recap along with recommendations on future areas of research.

3.2. Theoretical framework

3.2.1. Social proximity as joint experience in failed project applications

In collaborative knowledge networks, social proximity or relational embeddedness between a pair of actors can serve as a strong driver of tie formation (Ranjay Gulati & Gargiulo, 1999; Walker et al., 1997). This is especially relevant in a regional context, where actors are more likely to know each other and previous interactions are common. In general, social proximity is associated with the presence of some degree of familiarity and trust between two organizations, based on friendship, kinship or past experience (Boschma, 2005; Boschma & Frenken, 2010). In other words, if two organizations have a shared experience, they are more likely to engage in a future collaboration as opposed to two organizations that have never interacted with each other before. Repeated engagements, in the form of past partnership or R&D collaboration, produce stronger network ties (Granovetter, 1973) and often engender knowledge-based trust and resource sharing between participating entities (Bstieler et al., 2017; R. Gulati, 1995; Santoro & Saparito, 2003; Zaheer et al., 1998). Recurrent interactions can also provide a certain level of predictability that reduces the perceived risk of conflict (Usai et al., 2017; Uzzi, 1996). Evidence on the impact of social embeddedness for tie formation has been detected in EU-FP networks (Autant-Bernard et al., 2007; Paier & Scherngell, 2011), as well as national and regional R&D networks (Caloffi et al., 2015; D'Este et al., 2013; Heringa et al., 2016).

In the vast majority of empirical studies, social proximity has been operationalized in terms of prior *collaborative* experience. Yet, in publicly subsidized R&D networks, many partnerships never transition into project execution, as they fail to secure the necessary funding. Premature failures are also common across university-industry interactions, as actors may undergo a lengthy process of informal negotiations without ever reaching the point of a formal collaboration agreement, which leaves little evidence of their engagement. This means that many organizations acquire joint experience in project preparation and project application without necessarily gaining joint collaborative experience. Both types of engagement increase the social proximity between parties, but they also differ in at least two ways:

(1) Joint experience in failed project applications can provide only a partial insight into the credibility, reliability and knowledge repository of a partner organization. Actual collaborative experience in one or multiple projects is needed to build trust between parties, to assess their skill base and their willingness to share tacit knowledge (Gulati & Gargiulo, 1999). According to Das & Teng (2002) trust is built through a long process of indirect reciprocal exchanges, rather than short-lived interactions. In addition, only joint collaborative experience can highlight partners' attitude toward the management of sensitive IP rights in a way that prevents opportunistic behavior. In R&D projects, issues of commercialization are likely to surface in the final stages of a collaboration, or even after it has been completed.

    In that sense, we can argue that collaborative experience facilitates behavioral learning, whereby organizations learn about each other's' trustworthiness, routines and knowledge base by interacting continuously and performing tasks collectively over an extended period of time (Doz, 1996). The sheer preparation of a joint application does not allow for such behavioral learning to take place. Nevertheless, joint experience in failed project applications can also be demanding and time-consuming (Hünermund et al., 2022), and still offer insights into the overall *commitment* of a partner organization, and its capacity to *communicate* and *set shared objectives* during the initial stage of project preparation. We discuss those in further detail below.

(2) The second difference concerns the negative connotation of joint experience in failed project applications. Although it is often difficult to pinpoint the precise reasons for a consortium's failure (proposal quality, misalignment with policy priorities, noncompliance with evaluation criteria, or other), the rejection itself can send a negative *external* signal regarding the relational dynamics of partner organizations.

When the datasets used in empirical studies include only financed projects, researchers are unable to distinguish between applicants who have partnered in the preparation of an unsuccessful proposal and are thus acquainted with each other, and applicants who have perhaps not interacted in the past. Furthermore, it is difficult to assess how the rejection itself influences actors' collaborative behavior: whether it reinforces existing patterns of interorganizational coupling or it rather reverses them. We draw on the literature on social networks, social capital theory, and transaction cost theory to delineate how *joint* *experience in failed project applications* – an important but understudied facet of social proximity – is associated with tie formation between partners in R&D networks.

From the perspective of transaction cost theory, the organizational structures needed to set up an R&D alliance are costly (Amoroso, 2014). This is especially relevant when the actors involved operate in distinct sectoral and institutional settings. Some level of alignment between different norms, policies and strategies is therefore necessary to overcome divergent orientations (public vs. private entities) and knowledge appropriation (Al-Tabbaa & Ankrah, 2016; Ankrah et al., 2013; Bruneel et al., 2010; Muscio & Vallanti, 2014). The "hidden costs" associated with partner searching, negotiation and mobilization of administrative capabilities to coordinate communication, accrue even before the actual collaboration takes place. Furthermore, the process of preparing a joint R&D proposal implies an additional degree of organizational learning, as partners need to abide by external rules, typically set by the funding entity. Thus, in cases of rejected project applications, organizations may try and recover some of these hidden operating costs by re-submitting their proposal in later open calls. In such instances, the joint experience in failed project applications will serve as a driver of future tie formation. Evidence on the potential for capitalizing on the initial investment can be found in studies, which show a positive correlation between past partnership experience and lower application costs (Takalo et al., 2013). Furthermore, in a study on Spanish firms' participation in FP R&D consortia, Barajas & Huergo (2010) found that prior experience in proposals increases the probability of applying in the next edition, and more importantly – that the effect is higher if the previous proposal was rejected, providing some evidence on the presence of cost-optimization strategies among applicants.

At the same time, failure to obtain funding may also push organizations to search for substitute partners, contributing to a shift in existing collaborative patterns. The negative outcome of an application could reflect difficulties in overcoming relational dialectics in the initial project set-up. In certain situations, organizations may attribute the consortium's failure to other team members, which would make their joint application experience an unlikely predictor of future tie formation. While interorganizational trust and effective knowledge transfer require some joint collaborative experience (Bstieler et al., 2017), other crucial aspects of partner dynamics such as *commitment*, *communication* and *capacity to set shared objectives* can become apparent early on in the application stage (Rybnicek & Königsgruber, 2019; Smiljic, 2020). For instance, commitment refers to the level of dedication partners exhibit toward each other and the collective undertaking (Barnes et al., 2002). Individual levels of commitment will likely manifest in the preparation of a project proposal with some entities displaying a more proactive attitude than others. Similarly, difficulties in communication can also surface early on. In various case studies, experienced partners report struggles in achieving a common terminology, as team members may think they are talking about the same thing even if that is not the case in practice (Canhoto et al., 2016; Smiljic, 2020). This often generates misunderstandings, which could transpire in the final project proposal. Lastly, laying down clearly defined objectives and a shared vision for collaboration constitute an essential step in the alliance formation stage, since organizations will inevitably bring their own perspective and expectations and those could be challenging to reconcile (Barnes et al., 2002; Inkpen & Tsang, 2005). Spending too much time in exploratory ideation could compromise partners' ability to arrive at a common understanding of the *practical* needs and goals of the project (Canhoto et al., 2016). In that sense, joint experience in failed project applications can be enough to highlight tensions in the interaction between different organizations, prompting them to search for alternative partners.

Given that empirical evidence is inconclusive, and both scenarios appear plausible, we put forward two opposing hypotheses:

H1 (a): Joint experience in failed project applications is *positively associated* with future tie formation.

H2 (a): Joint experience in failed project applications is *negatively associated* with future tie formation.

3.2.2. The moderating effect of cognitive proximity

The literature on R&D networks has emphasized the role of cognitive proximity for the frictionless formation of interorganizational alliances. From a network perspective, cognitive proximity has been shown to positively impact tie formation (Broekel & Boschma, 2012; Cantner & Meder, 2007; Simensen & Abbasiharofteh, 2022; Werker et al., 2019), while management and innovation studies have provided a conceptual rationale for this relationship. Organizations need to possess sufficient absorptive capacity to be able to benefit from their interactions, that is: to identify, interpret and exploit knowledge embedded in partner entities (Cohen & Levinthal, 1990; Nooteboom, 2000). Some level of cognitive proximity is therefore required for organizations to reach mutual understanding and tap into the value of their partner's knowledge. At the same time, cognitive distance can also be viewed as an asset in terms of combining distinct knowledge bases. Empirical studies have shown that collaborations between dissimilar partners who bring diverse competencies to the table, are also associated with higher output, both in terms of publications (Werker et al., 2019), and number of explorative patents (Gilsing et al., 2008). The key argument here is that as cognitive distance increases, so will the opportunities for learning and novel recombination, but beyond a certain threshold, cognitive distance will begin to preclude mutual understanding, leaving partners incapable of utilizing these opportunities (Nooteboom et al., 2007; Wuyts et al., 2005). In sum, organizations may struggle to forge network partnerships with cognitively distant partners and will generally avoid it, but once a link has been established, the potential benefit of this relationship in terms of innovative output is likely to be significant (Werker et al., 2019).

In our case, given the difficulties associated with setting up a formal R&D partnership between cognitively distant partners, we can reasonably assume that organizations which have already "paid" the hidden operational costs of establishing a diverse consortium will be *particularly* motivated to capitalize on their initial investment, if their original project application happens to be rejected. In addition, the high reward associated with bringing on board cognitively heterogeneous team members is likely to further incentivize partners to maintain their relationship and re-apply together in subsequent open calls, once the initial barriers to establishing contact have been overcome. From the discussion above, it follows that if joint experience in failed project applications is indeed positively associated with future tie formation (H1 (a) holds true), the cognitive distance between partners will only strengthen this relationship.

H1 (b): If the relationship between joint experience in failed project applications and future tie formation is positive, the cognitive distance between partners will strengthen it.

Following the same logic with regards to the incentives to capitalize on connections with cognitively distant partners, if the relationship between joint experience in failed projects applications and future tie formation is negative, then we would expect that cognitive distance between partners should reduce the disposition to search for substitute partners. In other words, organizations may be more likely to replace partners with whom establishing a link is less costly, but not those cognitively distant ones with whom forging a connection required significant investment of time and resources.

H2 (b): If the relationship between joint experience in failed project applications and future tie formation is negative, the cognitive distance between partners will weaken it.

3.3. Methodological approach

To tackle the research questions presented above, we rely on the entire dataset of both approved and rejected R&D projects in the Valencian region. Recall that the dataset, as described in chapter 2, encompasses a total of 444 projects spanning a 7-year timeframe, and involving 554 distinct organizations. Here, we introduce our empirical approach, which includes the operationalization of critical variables and the selection of an appropriate econometric method.

3.3.1. Variables and models

Given the nature of our hypotheses, we seek to model the selection of R&D partners as a function of prior application experience while controlling for other well-known drivers of tie formation. Since the number of project applications is uneven across the 7 years of observation and AVI was only established in 2018, we decided to group the data into two periods: 2016-2020 (period t) and 2021-2022 (period t+1). This allows us to examine the structural properties of the regional knowledge network during the second period in light of the "history" of prior engagement in the first period, and it also ensures we have a balanced mix of R&D partnerships from both public funding agencies. Furthermore, the event we study is highly infrequent, so observing every potential dyad on an annual basis would significantly increase the share of 0s, making it harder to discriminate between the two scenarios (Diestre & Rajagopalan, 2012).

The unit of analysis is the dyad. We first construct a list of all possible partnership constellations between the 554 organizations present in our 7-year time window. In line with other studies, we assume dyads to be at equal risk (*a priori*) of submitting a joint application (Diestre & Rajagopalan, 2012; Rothaermel & Boeker, 2008). We treat links between nodes as bilateral (that is: A-B is identical to B-A). In this empirical set-up, partnership, be it successful or unsuccessful, is a rare event.

Our primary dependent variable ***Partner*** is a binary one, and takes the value of 1 if the two organizations in the dyad applied for collaborative R&D funding together as part of the same team in the second period t+1 (regardless of whether or not that application was successful) and 0 otherwise. By considering both funded and unfunded projects, we are able to fully capture the partner selection behavior of organizations rather than the selection behavior of the funding body.

Our primary independent variable ***FailedExp***, is also dichotomous and it is equal to 1 when the two entities in the dyad had only joint experience in failed project applications during the first period t. We also introduce a second dichotomous variable ***SuccessExp*** which is equal to 1 if the two entities had only joint collaborative experience (through successfully funded projects) during the first period. Making a clear distinction on the nature of actors' prior engagement allows us to capture the two facets of social proximity: one based on successful project application which results in joint project implementation, and one related solely to the joint experience of setting up a team and a project proposal, which does not materialize in a real collaboration. Since some of the dyads in the sample experience a mix of both scenarios, we include a variable ***MixedExp*** which is equal to 1 when the two organizations have had both successful and unsuccessful applications together at time t. For comparison purposes, we decided to elaborate a forth variable called ***CollabExp,*** which reflects the current standard for operationalizing past experience, as it is typically adopted in empirical studies. This variable takes the value of 1 when the two organizations in the dyad had *at least* 1 funded project in period t, that is: at least 1 prior collaboration, and 0 otherwise. In that sense ***CollabExp*** is "blind" to the presence of mixed or failed partnership experience.

We construct several control variables to ensure the robustness of the empirical analysis. First, we account for the geographical distance between organizations in the dyad by adding a variable called ***GeoDist***. There is now substantial empirical evidence indicating that organizations located closer to each other are, on average, more likely to collaborate (Howells, 2002; Katz, 1994; Ponds et al., 2007). ***GeoDist*** is continuous and it measures the straight-line distance (in km) between the registered addresses of the entities in the pair.

Furthermore, we control for the difference in project-related experience between the two organizations. In the context of competitive publicly-subsidized R&D schemes, the sheer knowledge and management capabilities accumulated from participating in open calls and preparing applications, be they successful or unsuccessful, is widely regarded as an important non-transferrable organization-specific asset (Cantner & Meder, 2007; Paier & Scherngell, 2011). To get a measure of ***ExpDiff***, we first calculate the cumulative sum of past applications for each of the two nodes, both successful and unsuccessful, in the first period. Once we have the sum for each node, we compute the arithmetic difference between the two values. Hence, ***ExpDiff*** is a continuous variable which reflects the difference in the number of past applications for each of the two organizations. A high value of ***ExpDiff*** implies greater discrepancy in the experience levels of the two partners, which means a more disassortative tie.

We also control for transitivity, which denotes a particular phenomenon in social networks, whereby two unconnected nodes with a mutual acquaintance tend to establish a link with one another, resulting in triadic closure (Ter Wal, 2014). Having a third-party reference seems to reduce uncertainty and asymmetric information, while also deterring future partners from exhibiting opportunistic behavior (Balland et al., 2016; Uzzi, 1996). Since triadic closure seems particularly relevant in a regional context, where the reputational consequences of violating established norms are greater (Gulati, 1995), we introduce a variable ***TC,*** which is equal to 1 if the two nodes in the dyad were connected by a third party in the first period through successful or unsuccessful partnership, and 0 otherwise.

We further control for the size of firms in the dyad. Large companies generally have more resources to engage in and maintain multiple R&D collaborations at a time, and they have incentives to join alliances as a way of maximizing spillovers and indirectly monitoring innovative activity in their field (Hernán et al., 2003). We introduce a dummy variable ***LargeFirm*** equal to 1 if at least one of the organizations in the dyad is considered a "Large enterprise", based on annual turnover.

Finally, since we are dealing with public open calls, we seek to control for specific requirements embedded in each program that may produce a disproportionately higher number of identical pairs, for instance: firm-firm, research center-firm, etc. For this reason, we introduce an additional dichotomized control variable (***PRO***) equal to 1 when at least one of the two entities falls in the category of university, (health) research institute, or technological institute.

Hence, our first model takes the following form:

$$Partner_{\mathrm{t+1}} = FailedExp_t + SuccessExp_t + MixedExp_t + ExpDiff_t + TC_t + GeoDist + LargeFirm + PRO$$

As a second step, we want to examine the moderating effect of cognitive proximity on the relationship between joint experience in failed project applications and tie formation. To do so, we have to limit our sample to dyads containing only private for-profits (approximately 81% of the original sample), as their NACE classification is known. We introduce a dummy variable called ***CogDist***, which reflects the sectoral heterogeneity between a pair of firms, based on their NACE codes. Its operationalization is relatively straight-forward: a value of 1 indicates that the two organizations operate in two completely different two-digit NACE Rev. 2 sectors. Using the degree of sectoral overlap as a proxy of cognitive proximity has been applied in other studies (Balland et al., 2016) and is consistent with the literature on related variety (Frenken et al., 2007) and the assumption that firms operating in similar sectors will likely possess related knowledge bases.

Since we have data on the age of all companies in the sample, we can introduce an additional control variable. By virtue of their accumulated industrial experience, mature companies may enjoy a certain status in their surroundings (Balland et al., 2016), prompting other companies to gravitate towards them for the formation of R&D partnerships. Furthermore, in line with the resource-based view, mature companies may find young start-up-like enterprises beneficial for acquiring complementary capabilities, which are missing or deficient internally (Ahuja, 2000a). To capture this effect, we introduce a continuous variable ***AgeDiff*** which reflects the difference in industrial experience between the two firms in the pair, as proxied by their years in operation. For the most part, we expect larger values of ***AgeDiff*** to represent an asymmetric partnership between a startup and a mature company. In this second stage, we omit the ***PRO*** variable used in the first model, as it is no longer applicable. All observations in the second model are firm-firm dyads.

$$Partner_{\mathrm{t+1}} = FailedExp_{\mathrm{t}} + (FailedExp_{\mathrm{t}}) * (CogDist) + CogDist + SuccessExp_{\mathrm{t}} + MixedExp_{\mathrm{t}} + ExpDiff_t + TC_t + AgeDiff + GeoDist + LargeFirm$$

Table 3.1 provides an overview of all variables used in the two stages of analysis.

Table 3.1. Descriptive statistics for the full set of variables.

Variable	Description	Obs.	Min	Max	Mean	Share of 0s
Dependent variable						
Partner	Dummy variable taking value 1 when the two organizations applied for funding together in the second period.	153181	0	1	0.006	99.0%
Explanatory variables						
FailedExp	Dummy variable taking value 1 when the two organizations have only joint experience in failed project applications in the first period.	153181	0	1	0.002	99.4%
SuccessExp	Dummy variable taking value 1 when the two organizations have only joint experience in successful project applications the first period.	153181	0	1	0.001	99.5%
MixedExp	Dummy variable taking value 1 when the two organizations have both successful and unsuccessful applications in the first period.	153181	0	1	0.0001	99.98%
CollabExp	Dummy variable taking value 1 when the two organizations have at least one successful application in the first period.	153181	0	1	0.001	99.9%
Moderator						
CogDist	Dummy variable taking value 1 when the two organizations operate in different two-digit NACE sectors.	123753	0	1	0.953	4.71%
Controls						
GeoDist	Continuous variable reflecting the straight-line geographical distance (in km) between the two organizations.	153181	0	272	64.601	0.001%
ExpDiff	Continuous variable reflecting the difference in experience between the two organizations in the first period. It is equal to the difference in the total number of applications filed by each of the two entities, including both successful and unsuccessful ones.	153181	0	28	1.391	34.5%
AgeDiff	Continuous variable which reflects the difference in "industrial status" between the two firms as proxied by their years in operation.	123753	0	141	18.64	2.1%
TC (Triadic Closure)	Dummy variable taking value 1 when the two organizations have a partner in common in the first period.	153181	0	1	0.020	98.0%
LargeFirm	Dummy variable equal to 1 if at least one of the organizations in the dyad is classified as a "Large enterprise".	153181	0	1	0.240	75.7%
PRO	Dummy variable taking value 1 when at least one of the two entities falls in the category of university, (health) research institute, or technological institute.	153181	0	1	0.074	92.2%

As expected, the share of 0 values in our dependent and independent variables is relatively high. It is also worth noting that roughly 95% of firm-firm dyads are cognitively distant. The maximum geographical distance between a pair of entities is 272 km, while the biggest difference in terms of number of applications between two organizations is 28.

3.3.2. Method

Because the dependent variable is binary in nature, we estimate the probability of joint application using a logit model, which is consistent with prior studies (Caloffi et al., 2015; Diestre & Rajagopalan, 2012; Heringa et al., 2016; Reuer & Devarakonda, 2017).

3.4. Results and discussion

Table 3.3 presents the results of the logistic regression, used to assess the influence of different types of prior engagement on the probability of tie formation using the entire sample of observations (Model 1). In a step-wise approach, we first introduce a baseline Model 1.0 which includes only controls, and excludes any information on prior experience. Model 1.1 adds prior collaborative experience ***CollabExp*** – a variable that ignores any information on rejected applications, and is reflective of what most empirical studies relying exclusively on funded projects would encompass. Model 1.2 incorporates our primary variable of interest: ***FailedExp,*** while Model 1.3 includes all three possible scenarios of prior engagement: ***FailedExp, SuccessExp,*** and ***MixedExp***[14]***.***

A correlation matrix is shown in Table 3.2. ***CollabExp*** and ***SuccessExp*** are expectedly highly correlated, since ***CollabExp*** is composed of both ***SuccessExp*** and ***MixedExp.***

Table 3.2 Variables correlation matrix (full sample).

Parameter	CollabExp	SuccessExp	FailedExp	MixedExp	GeoDist	ExpDiff	TC	PRO	LargeFirm
Partner	**0.055**	**0.040**	**0.179**	**0.056**	**-0.038**	**0.207**	**0.035**	**0.119**	0.002
CollabExp		**0.953**	-0.002	**0.304**	**-0.021**	**0.065**	-0.005	**0.024**	**0.012**
SuccessExp			-0.002	0.000	**-0.019**	**0.058**	-0.005	**0.017**	**0.011**
FailedExp				-0.001	**-0.027**	**0.089**	-0.006	**0.060**	**-0.016**
MixedExp					**-0.009**	**0.032**	-0.002	**0.024**	0.003
GeoDist						**-0.060**	**-0.057**	**-0.049**	**-0.012**
ExpDiff							**0.127**	**0.361**	-0.004
TC								**0.125**	-0.001
PRO									**-0.073**

Note: Values in bold are statistically significant (p < 0.01).

[14] As a robustness check, Model 1 was also estimated using the Quadratic Assignment Procedure. Results are available in Appendix 3.A.

First, we comment on the results of our baseline model, which indicate that all control variables are statistically significant at the 1%, except for triadic closure ***TC***, which is significant at the 10%. ***GeoDist*** has an expectedly negative coefficient, which means that actors who are located further away from each other are less likely to enter an alliance together, whereas actors who have asymmetrical experience in project applications, as reflected by our third control variable ***ExpDiff,*** are more likely to do so. Triadic closure (***TC***) was also found to be a predictor of future tie formation, along with the presence of a PRO or a large company in the dyad. Model 1.1 demonstrates that, consistent with prior studies, the existence of at least one successful collaboration between two organizations in the past increases the probability of them entering an alliance at time t+1. Exponentiating the estimated log-odds coefficient for ***CollabExp***, we calculate that the presence of at least one case of prior successful collaboration between two partners increases the odds of them re-engaging in t+1 roughly 5 times. This result, as mentioned earlier, overlooks the potential presence of rejected applications between partners.

Moving on to Model 1.2, we observe a positive and significant coefficient for our primary variable of interest, ***FailedExp.*** This suggests that prior joint experience in rejected project applications, similar to prior experience in successful applications, increases the probability of future tie formation. Finally, with respect to Model 1.3, we make two important observations. First, when controlling for all possible scenarios of prior engagement, joint experience in failed project applications seems to exert a positive influence on future tie formation, and a stronger one than joint experience in successful applications. Using odds ratio, we can say that when two organizations have partnered only in awarded projects, the odds of them re-engaging at time t+1 increase approximately 5 times, but if their experience is only one of failure, the odds of them re-engaging increase roughly 30 times. The difference between the estimated coefficients for ***SuccessExp*** and ***FailedExp*** was found to be statistically significant (based on a Wald test, p-value = 1.8e-8). This provides strong support for our first hypothesis (H1.a). Second, when breaking down the variable ***CollabExp*** into its two primary components: ***SuccessExp*** and ***MixedExp,*** we observe that the estimated coefficient for ***MixedExp*** is larger than that of ***SuccessExp*** and the difference is statistically significant (based on a Wald test, p-value = 0.027), meaning that the effect detected in many empirical studies, which rely exclusively on funded projects, may have been boosted by unobserved failures accompanying the successes.

Table 3.3. Results of the standard logit regression, assessing the probability of R&D alliance formation, using the full sample.

	Dependent variable: $Partner_{t+1}$			
	(Model 1.0)	(Model 1.1)	(Model 1.2)	(Model 1.3)
CollabExp		1.677^{***}		
		(0.258)		
FailedExp			3.410^{***}	3.444^{***}
			(0.161)	(0.160)
SuccessExp				1.601^{***}
				(0.291)
MixedExp				3.019^{***}
				(0.579)
GeoDist	-0.008^{***}	-0.008^{***}	-0.008^{***}	-0.007^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
ExpDiff	0.127^{***}	0.125^{***}	0.123^{***}	0.121^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
TC	0.239^{*}	0.280^{**}	0.417^{***}	0.468^{***}
	(0.132)	(0.132)	(0.132)	(0.132)
PRO	1.475^{***}	1.468^{***}	1.398^{***}	1.381^{***}
	(0.084)	(0.084)	(0.085)	(0.086)
LargeFirm	0.341^{***}	0.329^{***}	0.413^{***}	0.400^{***}
	(0.079)	(0.079)	(0.080)	(0.080)
Constant	-5.473^{***}	-5.482^{***}	-5.581^{***}	-5.593^{***}
	(0.064)	(0.064)	(0.066)	(0.066)
Observations	153,181	153,181	153,181	153,181
Log Likelihood	-4,770.546	-4,753.483	-4,603.504	-4,580.465
McKelvey & Zavorina R^2	0.155	0.154	0.152	0.150

Note: Robust standard errors are in brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

These findings have several implications. First, it appears that the importance of relational embeddedness or social proximity may, in fact, be underrated. While extant literature has illustrated the organizational propensity to replicate ties with former collaborative partners, our study suggests that this type of retention mechanism is at play even when the two entities have only joint experience in failed project applications. The organizations in our sample appear more inclined toward selecting partners with whom they have *some form* of shared experience, be it successful or unsuccessful, as opposed to a completely different organization, with whom they have no level of social proximity, at least not in the case of the three competitive programs we analyze. Thus, our study disputes existing assumptions on what type of prior interaction is necessary to stir further engagement and empirically demonstrates that the social proximity accumulated during a failed project application is in fact a sufficient condition for renewed partnership.

Second, our results provide strong evidence on the presence of cost optimization strategies among cooperating organizations. As argued earlier, the identification of suitable R&D partners and the establishment of a working relationship requires considerable transaction costs, which would be lost or "sunk" if the relationship is abandoned due to lack of funding. Naturally, partners will try to benefit from their initial investment by re-applying in subsequent open calls. In an earlier study comparing approved and rejected single-beneficiary projects, Barajas & Huergo (2010) found evidence of this pattern. They showcase how prior experience in FP proposals increases the probability of applying in the next edition and that the effect is higher if the previous proposal was rejected. Our results, however, build on and expand this argument: not only do organizations reapply when rejected, but they appear to do so with the same partners. In other words, they seek to recover not only the cost of elaborating a competitive project proposal, but also the search cost of finding appropriate partners, negotiating and consolidating a working relationship with them. The latter one may in fact represent a more substantial upfront investment of time and resources, than the former.

Third, the strong influence of joint experience in failed project applications signals that studies relying exclusively on financed R&D projects may, in fact, be missing an important antecedent of tie formation in the analysis of interorganizational networks. It was surprising for us that the effect of past collaborative experience, which has been well-documented in the literature as a strong driver of tie formation, was overshadowed by the effect of joint experience in failed applications. We considered the possibility that re-applying after rejection may occur faster, since implementing an actual R&D project is likely to take up a few years. Yet in our sample, the average project duration is only 2 years, and neither of the policy programs we considered explicitly prohibits funded actors from reapplying in the following year, so for those reasons we do not apply a time lag. Hence, our results suggest, rather crucially, that in the absence of information on rejected projects, empirical studies may end up overestimating the effect of prior collaborative experience.

Next, we move to the results of the second model, where we tested the moderating effect of cognitive proximity on the relationship between joint application experience and tie formation by using a subsample of firm-firm dyads. Table 3.4 summarizes the results, while the correlation between relevant variables is available in Appendix 3.B.

Looking at Model 2.0, once again we observe that ***FailedExp*** is positively associated with tie formation, and the effect appears to be greater than that for ***SuccessExp***. In both models 2.0 and 2.1, the cognitive distance between firms is negatively associated with tie formation. Exponentiating the coefficient for ***CogDist,*** we find that having two cognitively distant firms in the dyad decreases the odds of them collaborating roughly 3 times. However, the interaction term between joint experience in failed project applications and cognitive distance has a positive and significant coefficient. Prior successful collaborations (***SuccessExp***) remain positively associated with the probability of further engagement. When using a subsample of firm-firm pairs, the variable ***MixedExp*** is no longer a significant predictor of tie formation, and the observed standard errors are quite large, most likely because situations of mixed experience are extremely rare in the subsample we are focusing on. It is further worth noting that two of the control variables (***ExpDiff*** and ***TC***) used in the previous model 1 seem to lose their significance in this set-up, while the **AgeDiff** control variable we introduced has little effect on the probability of tie formation, and its coefficient is only significant at 10%.

Table 3.4. Results of the standard logit regression, assessing the probability of R&D alliance formation, using a subsample of firm-firm dyads.

	Dependent variable: *Partner*$_{t+1}$	
	(Model 2.0)	(Model 2.1)
FailedExp	4.051^{***}	1.577
	(0.285)	(1.037)
CogDist	-1.253^{***}	-1.423^{***}
	(0.166)	(0.164)
(FailedExp) (CogDist)		3.080^{***}
		(1.076)
SuccessExp	2.117^{***}	2.080^{***}
	(0.520)	(0.521)
MixedExp	-7.679	-7.652
	(199.245)	(199.176)
GeoDist	-0.010^{***}	-0.010^{***}
	(0.001)	(0.001)
ExpDiff	-0.002	-0.006
	(0.051)	(0.051)
TC	-0.418	-0.432
	(0.583)	(0.583)
LargeFirm	0.701^{***}	0.707^{***}
	(0.125)	(0.125)
AgeDiff	0.005^{*}	0.005^{*}
	(0.003)	(0.003)
Constant	-4.767^{***}	-4.624^{***}
	(0.175)	(0.171)
Observations	123,753	123,753
Log Likelihood	-1,917.409	-1,909.654
McKelvey & Zavorina R^2	0.130	0.135

Note: Robust standard errors are in brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Next, we interpret these results. First, the negative coefficient of ***CogDist*** is in line with the theory on transaction cost economics. Cognitive distance increases the cost of establishing R&D partnerships, since organizations have to overcome significant barriers in understanding and interpreting knowledge, routines and practices of their potential partner. As expected, organizations are generally unwilling to embark on partnerships that entail a high cognitive distance. However, the positive coefficient of the interaction term confirms our hypothesis H1.b. Since interpreting interaction effects is difficult in nonlinear models including logistic regressions (Hoetker, 2007; Murphy & Aguinis, 2022), we examine the effect graphically. Figure 3.1 displays the interaction at 90% confidence intervals. It appears that once two cognitively distant firms have invested in establishing a partnership, they are far more likely to try and benefit from their initial investment when faced with a rejected project proposal, than cognitively proximate firms. The relationship of the former, when first established, proves particularly "sticky".

The coefficients of ***ExpDiff*** and ***TC*** are not entirely surprising, given that the difference in collaborative experience is more relevant when the sample includes a wider variety of institutional partners (as corroborated by the correlation coefficient between ***PRO*** and ***ExpDiff,*** see Table 3.2). PROs, by virtue of their mission, human and technical resources, can afford to engage in multiple alliances at a time, as opposed to private companies which generally find it difficult to accrue such levels of collaborative experience. As for **AgeDiff**, we note that despite our initial expectations that startups and mature companies might gravitate toward each other in search of complementary assets, we observe no such signs of disassortative behavior among the studied firms.

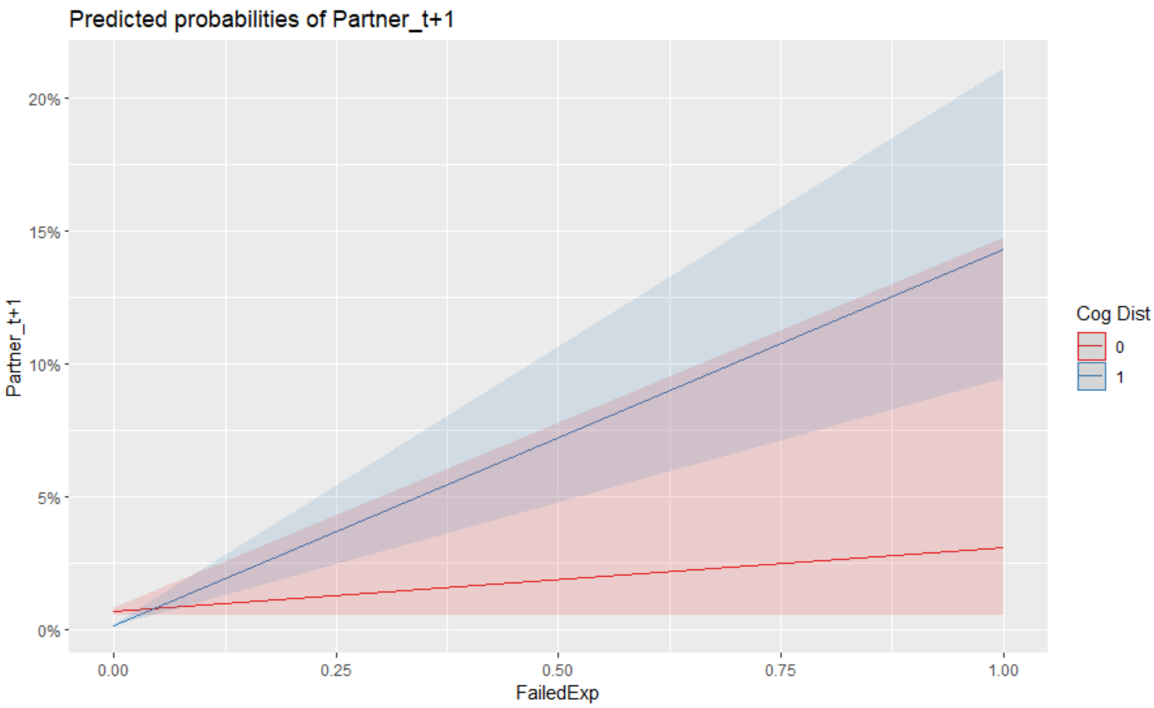

Figure 3.1. Interaction effect between joint experience in failed project applications and cognitive distance between partners.

3.5. Conclusion

This chapter aimed to investigate empirically how joint experience in rejected project applications – an important but overlooked facet of relational embeddedness – influences tie formation in the context of policy-supported R&D networks, and how this relationship plays out for partners with different levels of cognitive proximity. Considering the prevalence of rejection in competitive R&D schemes and the lack of studies assessing its impact on organizations' collaborative behavior, this study delivered several important insights.

First, it added to existing literature on social proximity or relational embeddedness, by conceptualizing a different form of prior engagement based on joint experience in failed project applications. Thus, it responded to recent calls for adopting a more comprehensive, portfolio view of interorganizational networking (Balland et al., 2022), by considering the full blend of possible interactions, be they successful, unsuccessful or mixed. The study provided empirical evidence that rejection is indeed a strong driver of future tie formation. Moreover, it showed that organizations are generally reluctant to abandon already established work relationships that required significant search costs, even if funding was not granted for the execution of a joint project. It appears that the stickiness of this relationship between rejected partners is, in fact, stronger when they are cognitively distant. Second, we showcased that in the absence of information on failed project applications, researchers may struggle to disentangle the influence of prior collaborations, since some of the observed effect may be boosted by unobserved failures accompanying the successes. Further research is needed to determine whether this observation holds true across different types of publicly-funded R&D networks.

Finally, the results of this study can be particularly beneficial for policymakers, as the introduction of competitive R&D schemes to stimulate collaboration between heterogeneous actors has been a key go-to instrument for many regional and national governments, not just in Valencia. Our analysis of three regional programs highlighted the important role these instruments play in generating a "hidden" network of partners, whose experience in preparing a joint application, even when rejected, serves as a strong motivation for further re-engagement. The observed propensity toward cost optimization implies that local actors will pursue alternative ways to finance their project plans, either by re-applying to the same program or to a different national or even international R&D scheme. In that sense, the social proximity generated by government-run programs is hardly an "expendable" asset we can expect to disappear after a project application has been rejected. Rather, it persists as an unobserved factor driving the evolution of interorganizational relations over time.

It is important to acknowledge the limitations of our data. First, in modelling future tie formation, we relied exclusively on formal R&D project applications, even though in practice local actors may be familiar with each other through alternative programs or schemes. A second limitation concerns the typology of the projects used. We relied on open calls from two independent institutions, which have unique requirements that may influence the organizational composition of the respective projects. While we sought to control for those differences in the empirical model, it is nonetheless important to keep them in mind when interpreting the results.

Future research can complement these gaps by incorporating primary data into the analysis. This means supplementing the observed collaborative behavior of regional actors with their self-reported motivations for picking one partner over another. Of course, this implies surveying or interviewing a sample of the regional actors, but it could provide highly valuable insights that allow policymakers to better predict and manage the evolution of the regional knowledge network.

Chapter 4. Challenging assumptions on strong ties in R&D networks

This chapter is based on the paper:

Yankova, D., D'Este, P. & García-Melón, M. (2023) "Does repetition equal more of the same? Challenging assumptions on strong ties in R&D networks". DRUID Working paper [Submitted to "PLOS ONE" journal; status: under review]

Abstract

Despite organizations' documented tendency to repeat interactions with prior partners, scholarly understanding on the implications of recurring collaborations has been fairly limited. This study investigates whether and under what conditions organizations use repeated engagements to explore new topics (what we call diversification), as opposed to deepening their expertise in a single one (specialization). The empirical analysis is based on the Spanish region of Valencia and its publicly-funded R&D network. Employing lexical similarity to compare the topic and content of project abstracts, we find that strong ties are not always associated with specialization. Yet, diversification is more likely when at least one of the partners mobilizes a growing network of diverse contacts and can access novel knowledge.

4.1. Introduction

The literature on inter-organizational networks has demonstrated organizations' proclivity to repeat interactions with prior partners, resulting in stronger ties and a reinforcement of existing network structures (Ranjay Gulati & Gargiulo, 1999; Walker et al., 1997). Empirical studies have documented this type of organizational inertia in partner selection within both national and international R&D networks (Autant-Bernard et al., 2007; Caloffi et al., 2015; D'Este et al., 2013; Heringa et al., 2016; Paier & Scherngell, 2011). Yet, the implications of repeated engagements for knowledge transfer remain an issue of contested debate. Do organizations leverage repeated ties to build expertise and specialize in a particular thematic domain or could strong links also be associated with the exploration of different topics? The answer to these questions can shed light on the value of repeated ties for individual and collective performance.

Innovation scholars have argued that strong bonds between partners are subject to declining marginal benefits (Gargiulo & Benassi, 2000; Masciarelli et al., 2010). At first organizations accumulate gains from solidifying existing relationships, as transaction costs decrease (Dyer & Singh, 2011) and the transfer of complex and tacit knowledge becomes easier, but beyond a certain threshold the learning potential for both parties may be exhausted (Becerra et al., 2008; Dyer & Singh, 2011; Hansen, 1999; Sorenson et al., 2006). Social embeddedness can begin to act as a filter for the entry of new knowledge and ideas, causing cognitive isolation and suboptimal innovative performance (Goerzen, 2007; Uzzi, 1996; Uzzi & Spiro, 2005).

So far scholarly understanding on the role and functionality of repeated interactions has been constructed independently and with little consideration for the nature of ties (Reagans & McEvily, 2003; Snijders, 1999). Few studies have looked explicitly at interaction processes and the strategic decisions partners make when repeating collaborations (Madhavan & Prescott, 2017). The goal of this chapter is to shed light precisely on this issue, by comparing the topics of recurring collaborations. It investigates *whether* and *under what conditions* organizations use repeated engagements to explore new topics (what we call diversification), as opposed to deepening their expertise in a single one.

This question is important for several reasons. If two organizations systematically exploit the same topic and tap into the same knowledge domain, their interactions will likely yield benefits at first, but hinder long-term innovative performance. If, however, subsequent collaborations begin to explore different topics, either because *a priori* the organizations involved possess a diverse internal repository of competencies and skills, or because they are capable of continuously sourcing novel knowledge through additional partnerships, the prospects of decreasing marginal benefits may weaken. Hence, the basic premise of this chapter is that the relationship between strong ties and performance will be at least partially contingent on the nature of the exchange between partners, and whether they choose to specialize or diversify in repeated engagements. For the rest of this document, we will use the term *specialization* to refer to consistent *exploitation* of the same topic in instances of repeated engagement, while *diversification* refers to the *exploration* of new topics, which differ from the one tackled in the first instance of engagement between partners.

Exploring the connection between specialization and strong bonds merits scholarly attention also as a departure from the structuralist perspective, which dominates knowledge network studies, and which treats inter-organizational ties as virtually homogeneous (Borgatti & Foster, 2003; Madhavan & Prescott, 2017; Phelps et al., 2012). By qualifying strong ties based on their nature and content, we can learn about the specific functions that seemingly identical types of relationships exercise in the context of inter-organizational networks (Owen-Smith & Powell, 2004; Podolny, 2001; Snijders, 1999).

To conduct the empirical analysis, we concentrate on the Spanish region of Valencia. We collected information on all R&D partnerships, formed between 2016 and 2022, which were awarded a public subsidy by one of the top two regional sources of innovation-related funding. The final dataset of 194 realized collaborative projects was used to map the local inter-organizational network and explore to what extent repeated engagements between partners in the 7 years of observations were associated with thematic specialization. We also test how partners' access to diverse knowledge and resources influenced the likelihood of them specializing or diversifying in subsequent collaborations. Given the rich literature on the benefits of degree centrality for learning, knowledge recombination and sustained innovative performance (Ahuja, 2000b; Graf & Krüger, 2011; Quintane & Carnabuci, 2016; Schilling & Phelps, 2007), well-connected actors may be more likely to explore new topics when re-engaging with the same partner.

This chapter adds to a growing stream of literature that recognizes the importance of strong ties as a frequent phenomenon in interpersonal and interorganizational networks (Dahlander & McFarland, 2013; Inoue & Liu, 2015; Tóth et al., 2021). It aims to illuminate the interplay between tie strength and specialization. The contribution is thus twofold: first, we develop a theoretical argument to suggest that the relationship between repeated engagement and specialization is fundamental for disentangling the effect of social cohesion on individual and collective performance. Though previous studies have demonstrated a clear link between network structural properties and actors' performance (Ahuja, 2000b; Baum et al., 2000; Giuliani & Bell, 2005; Morrison, 2008), there is still relatively little understanding on the precise mechanisms which underlie this relationship (Madhavan & Prescott, 2017; Phelps et al., 2012; Reagans & McEvily, 2003). Second, from a methodological perspective, our study applies recent advancements in machine learning and natural language processing (NLP) techniques to build a measure of specialization that is based on the lexical similarity between project abstracts. Instances of NLP usage in the innovation and management literature are increasing (Balsmeier et al., 2018; Kaplan & Vakili, 2015; Kelly et al., 2021), but they have concentrated primarily on patents' textual data, whereas our goal is to showcase the potential of such methods to advance scholarly understanding of R&D networks and the value of inter-organizational linkages.

From a policy standpoint, our analysis is also highly relevant. In regional R&D networks, specialization may be a desirable outcome if efforts are directed toward building competitive advantage in nascent or underexplored economic domains. Conversely, thematic specialization can be highly undesirable if the network is stagnating and policymakers are looking to branch out of existing development paths. Therefore, understanding how and under what conditions repeated engagement between regional partners is associated with thematic specialization could help policymakers steer more effective network interventions.

The chapter is structured as follows: section 4.2 lays the theoretical foundation of the study and relates it to the literature on inter-organizational networks and social capital. Section 4.3 introduces the characteristics of the dataset, and section 4.4 outlines our approach to operationalizing thematic specialization by building a measure of abstract similarity. Section 4.5 details the results of the analysis, while section 4.6 discusses their implications for theory and policy.

4.2. Theoretical background

4.2.1. The connection between repeated ties and specialization

A growing body of literature points to the importance of social embeddedness in driving the structural evolution of inter-organizational networks (Boschma, 2005). The formation of new partnerships between organizations is perceived in the context of their existing social structure and their history of prior ties (Ranjay Gulati & Gargiulo, 1999; Walker et al., 1997). Past engagement seems to impact the course of future cooperation in a path-dependent fashion, as former ties repeat themselves. This form of organizational inertia in partner selection has been observed in industry networks, cluster networks, regional and international R&D networks (Autant-Bernard et al., 2007; Balland et al., 2016; Caloffi et al., 2015; D'Este et al., 2013; Paier & Scherngell, 2011; Xavier Molina-Morales et al., 2015).

The implications of strong ties for performance have been the subject of many empirical studies. Some highlight the benefits of strong bonds for fine-grained knowledge sharing, in line with Coleman's theory on social capital (1988). Repeated engagements tend to engender "relational" trust between participating entities (R. Gulati, 1995; Santoro & Saparito, 2003; Zaheer et al., 1998). This can in turn reduce actors' perception of expected opportunistic behavior, decrease transaction costs and ease the transfer of both complex and tacit knowledge (Becerra et al., 2008; Dyer & Singh, 2011; Hansen, 1999; Sorenson et al., 2006). On the other hand, strong ties between partners may also reinforce retention mechanisms and prevent the inflow or nonredundant information (Gargiulo & Benassi, 1999; Uzzi, 1996; Uzzi & Spiro, 2005). When organizational partners become so narrowly focused on a particular type of activity, a transition toward new developments becomes difficult, leading companies to display inferior economic performance (Goerzen, 2007).

Taking on board both perspectives, scholars have settled the relationship between strong ties and performance as an inverted U-shape. Organizations benefit from consolidating strong relationships up to a certain level, beyond which social embeddedness can act as a filter for the entry of new knowledge and perspectives, causing cognitive isolation and suboptimal innovative performance (Gargiulo & Benassi, 2000; Masciarelli et al., 2010). This situation is also known as "the proximity paradox", since the same factors that drive actors to connect and exchange knowledge may also lead them to innovate less in the long run (Boschma & Frenken, 2010; Broekel & Boschma, 2012).

In this chapter, we argue that the consequences of repeated collaborations for individual and collective performance cannot be fully disentangled without examining the nature of ties, and acknowledging that organizations may leverage repeated interactions for different purposes. In his seminal work on the strength of weak ties, Granovetter (1973) noted that “treating only the strength of ties ignores […] all the important issues involving their content”, and stressed that the relationship between strength and degree of specialization of ties deserves further analysis. In addition, as pointed out by Reagans & McEvily many studies tend to infer knowledge transfer from the association between network structure – including tie strength – and performance, without directly examining the essence of the exchange (Ahuja et al., 2012; Madhavan & Prescott, 2017; Reagans & McEvily, 2003).

In order to unpack this association, we employ the so-called “connectionist” perspective (Borgatti & Foster, 2003), which looks beyond the structural or topological properties of the network, and treats ties as conduits of knowledge and resource flow (Owen-Smith & Powell, 2004; Podolny, 2001; Snijders, 1999). This approach acknowledges that seemingly identical types of links could exercise distinct functions and transmit varying kinds of resources. Furthermore, it recognizes that organizations in alliance networks are not simply “helpless targets of structural influence”, but active agents who make conscious decisions about the way they leverage strong bonds (Madhavan & Prescott, 2017).

Take for instance the following scenario: if two organizations systematically tackle the same topic in multiple collaborations, their interactions may follow the inverted U-shape scholars describe, whereby specialization would at first yield positive outcomes, but if continued for too long would hamper innovative performance. In the process of specializing itself, organizations are expected to tap into the same knowledge pool, building expertise at the start but eventually exhausting the recombination potential. If, however, subsequent collaborations begin to tackle different topics, either because a priori partners possess diverse internal repository of competencies, or because they are capable of sourcing those through third-party links, the graph of decreasing marginal benefits may take a different shape. At the very least, we can expect the threshold of redundancy, when the two partners have little learning space left, to become higher. Hence, the relationship between strong ties and performance is contingent on the nature of the exchange between partners, and whether or not they choose to specialize or diversify in repetitive engagements. This is not to suggest that specialization or diversification is inherently preferable. Rather, our goal is to illustrate that structurally equivalent relations, in the form of strong bonds, can have very different consequences for knowledge exchange and learning depending on the nature of the collaboration itself. To examine the heterogeneity of organizational approaches to repeated collaborations, we pose the following research question:

Q1: To what extent do organizations leverage repeated collaborations to exploit the same topic multiple times (specialization) or to explore new ones (diversification)?

4.2.2. Factors that moderate the relationship between repeated ties and specialization

The extent to which organizations use repeated collaborations to specialize in prior topics or diversify into new ones, may be influenced by their access to complementary knowledge and resources from third parties. Assuming that the knowledge repository of an entity is not static, forming alliances with a wide range of partners creates new pipelines for fresh ideas, perspectives, and information to flow (Owen-Smith & Powell, 2004; Podolny, 2001).

This may in turn inspire greater diversification in the topics and content of repeated collaborations.

So far, multiple empirical studies have demonstrated that the size of a firm's (ego) network, defined in terms of both direct and indirect contacts (alters), is positively associated with innovative output (Ahuja, 2000b; Baum et al., 2000; Schilling & Phelps, 2007). The theoretical framework, underlying these findings, assumes that well-connected organizations will have a more timely access to larger volumes of information through their established relationships. Yet, in line with the resource-based perspective (Lavie, 2006), some researchers have argued it is not the sheer number of connections that matters, as much as the diversity of knowledge which can be sourced through direct relationships (Baum et al., 2000; Stuart, 2000). In other words, a focus on the composition of the ego's first-order network, and more specifically the number of unique partners, may be more appropriate. Assuming that each organization holds a distinct set of assets and capabilities, direct relationships to multiple unique organizations may provide the best access to non-redundant knowledge and resources. In other words, we posit that the number of new, unique (i.e. non-shared) connections partners build before re-engaging with each other again may influence their propensity to explore new topics in a repeated exchange. To examine this issue further, we propose a second research question:

Q2: To what extent does partners' range of unique connections to other organizations inspire diversification in their repeated collaborations?

4.3. Data collection and overview

To investigate the relationship between repeated engagement and specialization, we refer once again on the dataset of policy-induced R&D project applications in Valencia, which we describe in detail in chapter 2. However, for the purposes of this study, we focus exclusively on the catalog of approved R&D applications, which represent *de facto* realized collaborations. According to public records, in the period 2016-2022, AVI and IVACE together funded a total of 220 collaborative R&D projects, under the three lines of action: "R&D in cooperation" (IVACE), "Strategic projects in cooperation" (AVI) and "Consolidation of the business value chain" (AVI).

Once the list of all 220 projects and their team members was compiled, a separate search was performed to collect textual descriptions for each project via several channels: (a) the official website of an organization involved in the project, (b) newspaper articles, or (c) the website of the funding entity. When no information about the collaboration was available online, we requested a brief description of activities from the principal investigator of the leading organization. Thus, our final sample consists of 194 R&D projects with a description longer than 50 words. This represents 88% of the entire list of funded projects in the time period of study. For the remaining 12% we were either unable to obtain a textual description or the one we had was too short and therefore insufficient to carry out a meaningful textual analysis. Most project descriptions mention the objective of the partnership, planned activities and expected results. As shown in Figure 4.1, abstract length varies between 50 and 450 words, with only a few exceptions of up to 750 words.

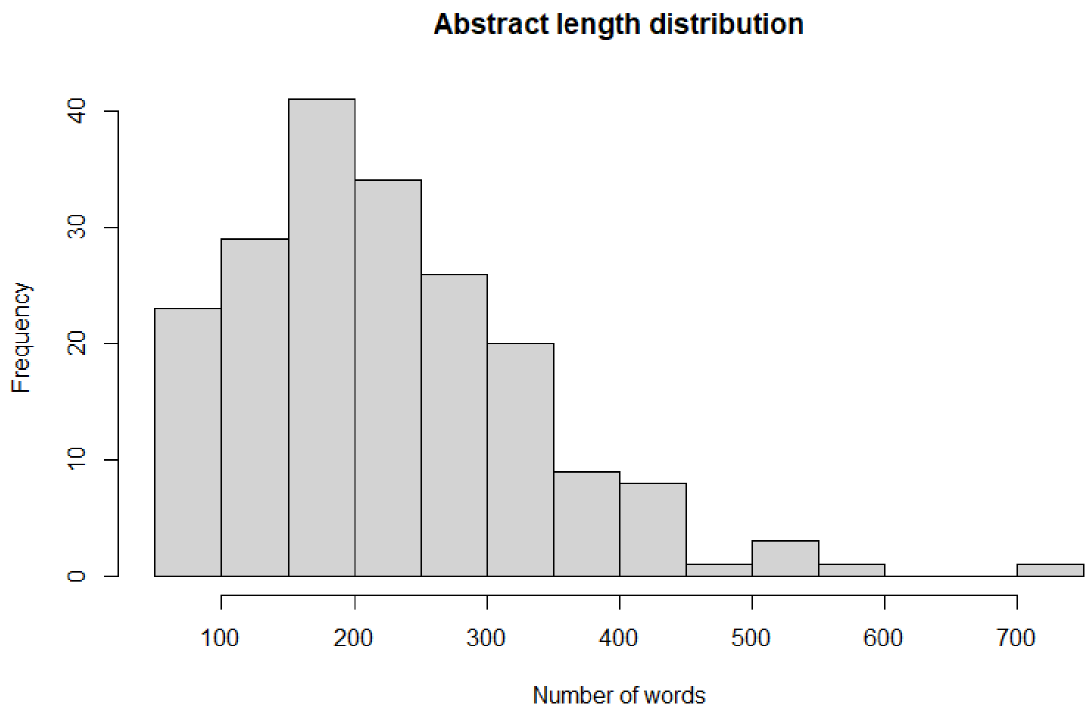

Figure 4.1. A histogram of project-abstract length, measured in word count.

The resulting R&D network consists of 362 individual organizations. 78% of them are private for-profits and about a third of all entities (nodes) participated in more than one project. The total number of realized links (edges) is 779, and roughly 5% of them were repeated at least once in the 7-year period. Table 4.1 provides descriptive statistics of the final sample, on which the analysis in this chapter was performed. It represents a sub-sample of the dataset we presented in chapter 2.

Table 4.1. Descriptive statistics of the final sample, used in this chapter.

Total number of collaborative projects	**194**
Min project size (number of partners)	2
Max project size	6
Average project size	3.2
Total number of organizations	**362**
Associations	7
Health research institutes	6
Private for-profits	282
Technological institutes	9
Research institutes	5
University departments and university-affiliated research centers	44
Others	9

It is important to note that in this chapter university-type beneficiaries were disaggregated into specific departments and teams. This means that every time a grant resolution referred to a particular university, a manual search was performed to identify the exact entity within the university structure that engaged in the collaboration. This allows us to build a fine-grained image of the regional R&D network and more importantly – it facilitates the operationalization of repeated engagements. Take for instance the following scenario: a company *x* completing two projects with university *y* can hardly be considered a case of repeated engagement, unless we can confirm that both instances concerned the same department or research team within the university (disaggregate level). More details on the operationalization of all variables is provided in the following section 4.4.

4.4. Variables and methods

To answer the main research questions, we build a 2-step approach and our unit of analysis is the pair of R&D projects. First, we construct all possible combinations of project pairs. Since our sample consists of n = 194 projects, all pairs amount to N = n*(n-1)/2, or 18721. Then we compare those that share a common dyad of partners – what we consider instances of repeated collaboration – to those that do not. We use descriptive analysis to shed light on the first research question. In the second stage, we isolate only project pairs which represent instances of repeated collaboration, meaning: they share at least one partner dyad in common (75 pairs in total), in order to test how the access to diverse sources of knowledge and resources of the two organizations influences the observed degree of specialization in their repeated engagements. In this second stage, we adopt an econometric approach and run a beta regression model, which is particularly suitable when the variable of interest is continuous and restricted to the interval (0,1) (Ferrari & Cribari-Neto, 2004). Below we elaborate the operationalization of our dependent and independent variables.

4.4.1. Constructing a measure of specialization

Since we are interested in analyzing whether organizations leverage repeated collaborations to specialize in a particular topic or to explore new ones (diversification), our primary dependent variable compares the thematic similarity between pairs of projects. Let us first revise the logic of this approach before diving into the empirical calculation.

Following the “connectionist” view of inter-organizational ties as pipelines that transmit tangible and intangible resources (Borgatti & Foster, 2003; Podolny, 2001), one arguably reliable way to infer the nature of those “pipe” flows is by tracing the description provided by the actors themselves. Joint project abstracts, or other types of descriptive project documentation, tend to provide sufficient information on – among other things – the specific area of intervention that partner organizations are focusing their collaboration on. Analyzing large volumes of text, however, is both challenging and burdensome. Fortunately, recent advancements in machine learning and NLP have opened up new possibilities for systematic interpretation of textual documents, including thematic classification and comparison.

Instances of NLP application in the innovation and management literature have proliferated, but so far they focus primarily on textual data from patents (Feng, 2020). Balsmeier et al. (2018), for example, introduce a measure of patent novelty, based on the first occurrence of a word in the patent corpus. Kaplan & Vakili (2015) use topic modelling, an unsupervised machine learning technique, to uncover the emergence of new topics in patent data and interpret those as cognitive breakthroughs. Also relying on textual analysis, Kelly et al. (2021), construct a measure of lexical similarity to quantify commonality in the topical content of patents, in order to identify significant ones – that is patents whose content is distinct from prior patents (more novel), but similar to future ones (more impactful).

Here we propose to leverage some of these advancements to measure the lexical similarity between pairs of R&D project abstracts so as to discern if repeated collaborations deal with the same topic. Lexical similarity is determined by the degree of lexical overlap, that is: how many terms from document i also appear in document j. It is a corpus-based method, which takes into account the co-occurrence of words across the entire collection of documents (corpus) (Chandrasekaran & Mago, 2021; Gomaa & Fahmy, 2013). We assume that projects whose descriptions show high levels of lexical similarity represent collaborative work in thematically proximate fields. When those projects were carried out by the same teams, we can interpret their repeated engagements as a continuation of previous work, or a form of specialization. Alternatively, lower similarity between project descriptions suggests that partners likely explored a completely different topic in their subsequent collaboration.

The following section details all steps in the calculation of the Abstract Similarity Score.

Measuring abstract similarity

Based on the sample of 194 regional collaborative R&D projects, we calculate a lexical text similarity score for all possible pairs of project abstracts (18721 pairs). We begin by constructing a document-term matrix (DTM), whereby each row represents a unique document (project abstract) and each column represents one term (word). The value of each matrix cell ij reflects the number of times term j appears in document i. Before creating the DTM, the corpus of abstracts is processed to remove punctuation, numbers and stopwords, such as pronouns, articles, specific verbs, and other common speech elements which carry little useful information. In addition, terms are trimmed to their stem without prefixes and suffixes, to avoid double-counting.

As highlighted by Kelly et al. (2021), a key consideration in building any similarity metric for a pair of text documents is to appropriately weigh the words by their importance. This is particularly crucial for our sample, since project descriptions follow a common structure and certain words (“objective”, “results”, “activities”) will be registered with greater frequency across the majority of text pairs, making them appear more similar than they really are. To account for that, we employ the “term-frequency-inverse-document-frequency” (TF-IDF) transformation method.

This method weights the registered occurrence of term j in document i, relative to its occurrence in the entire collection of documents Ni (Sammut & Webb, 2017). Consider the following equation:

$$W_{ij} = TF_{ij} * IDF_j = \left(\frac{n_{ij}}{l_j}\right) * \left(log\,\frac{N_j}{N_{ij}}\right)$$

where nij is the number of times word j appears in document i, li the length of i in terms of total number of words, Ni is the total number of documents in the corpus, while Nij is the number of documents in which term j appears. The terms with higher Wij will be those that appear relatively often within a document, but do not appear in the rest of the corpus.

These terms are therefore more representative of the document's semantic content. Put differently, the TF-IDF approach allows us to overweight words which are more diagnostic of an abstract´s topical content (Kelly et al., 2021).

The final DTM (dimensions: 194 x 3352) is quite sparse, and most term-frequency vectors contain many 0 values. To estimate how textually close two projects abstracts are, we use cosine similarity, which is measured by the cosine of the angle between a pair of term-frequency vectors and determines whether they are roughly pointing in the same direction. It is one of the earliest and most widely used distributional measures. The advantage of using cosine similarity is that it ignores zero-matches, essentially safeguarding against false positives. For example, two term-frequency vectors may have many 0 values in common, meaning that the corresponding documents share few words, but this does not make them similar. Cosine similarity focuses on the words two vectors have in common and the respective weight of these words (Han et al., 2012). It is a continuous metric that goes from 0 to 1. A high similarity score implies that two abstracts use the same set of words in the same proportion, while a lower similarity value shows no significant overlap between the texts.

Illustrating the method with practical examples

Next, we discuss the meaning of the Abstract Similarity Score in practice. We examine first a pair of projects, which has one of the highest similarity scores (0.44) in our sample. Given that abstract length surpasses 300 words, we have included only selected excerpts from the text, which highlight succinctly what each project is about.

Project **a** description

"[…] Detection and control of sulphate-reducing bacteria in drinking water infrastructures is presented in order to detect the critical points of the drinking water distribution network and implement the necessary improvements to reduce the risk of leaks and prevent water from losing its quality. The main objective of the project is to control and eliminate the development of sulphate-reducing bacteria in drinking water infrastructures through the development of new techniques for the detection of microorganisms and the functionalization of surfaces, reducing the risk of breaks and leaks and increasing the resilience of the drinking water distribution system."

Project **b** description

"[…] Optimization of the hydraulic performance of the drinking water network by means of optical fibre with the aim of detecting possible leaks generated in the supply network, as well as locating them throughout the system. The main objective of the project is the development of a system for detecting leaks and structural failures in drinking water pipes, accurate and economical, operating continuously, based on photonic technologies, and more specifically on "Distributed Acoustic Sensing" (DAS), which can be implemented in pipes in service and that its installation serves as a primary structural element for the implementation of future fiber optic sensors, specifically of water quality, without the need for new wiring."

From the excerpts we can see that the two projects rely on different technologies but the area of intervention is clearly similar: improving the resilience of the drinking water distribution system. Next, we compare a second pair of projects with a low similarity score (0.02), selected at random.

Project **a** description

"The main objective of this project is the research and development of an intelligent tool for dermatological exploration that assists in the detection and delimitation of the main types of skin cancer and does so in real time without the need for biopsy and through an automated and contactless technique […]"

Project **b** description

"The aim of this project is […] to improve the management of artificial wetlands for wastewater treatment, to naturalize their effluents, to minimize the impact on the receiving aquatic environment and to contribute to the mitigation of climate change […]"

Clearly, the two projects deal with two very distinct topics, and the algorithm is accurately assigning a low similarity score to this particular pair. It is worth emphasizing that the cosine similarity method itself does not tell us if the project partners utilize similar technologies, knowledge or resources, as such kind of information would be difficult to extract and requires more fine-grained and detailed textual data (for example: extended project reports). Nevertheless, we consider that the project abstracts we work with are sufficiently descriptive to allow for meaningful comparisons of thematic fields. They communicate the main area of intervention in a concise and straightforward fashion, while avoiding unnecessary noise-generating words. Therefore, we can discern if two projects represent a thematic extension of each other. While it is not a perfect measure, the abstract similarity score signals if organizations engaged in repeated collaborations continue deepening and specializing, broadly speaking, along the same topic.

4.4.2. Independent and control variables

In the first stage of the analysis, where we want to check the similarity of projects in cases of repeated collaborations, we introduce a dummy variable, called **SharedDyad**, which is equal to 1 if the two project teams have *at least 2* organizations in common. In other words, for a pair of project abstracts i_1-i_2, we compare the team of partners T_{i1} to the team of partners T_{i2}. Assume that project i_1 was carried out by T_{i1} = [A, B, C], while project i_2 was executed by T_{i2} = [A, B, D, E], where A, B, C, D and E are five unique organizations. Since the tie between A and B has persisted in both projects, the variable **SharedDyad** would assume the value of 1 even though the two teams are not completely identical and contain additional partners [E, C and D]. **SharedDyad** is equal to 0 when the two teams have only 1 or no partners in common. This approach is consistent with other studies on team repetition (Inoue, 2015). Figure 4.2 provides an illustrated example to further clarify the operationalization of **SharedDyad**. We opted for a dichotomous variable, rather than a categorical one, because in our sample instances of 3 shared partners were extremely rare.

Therefore, we cannot make a strong distinction between repeated dyads vs. repeated triads, but we believe that comparing the two cases may yield interesting insights.

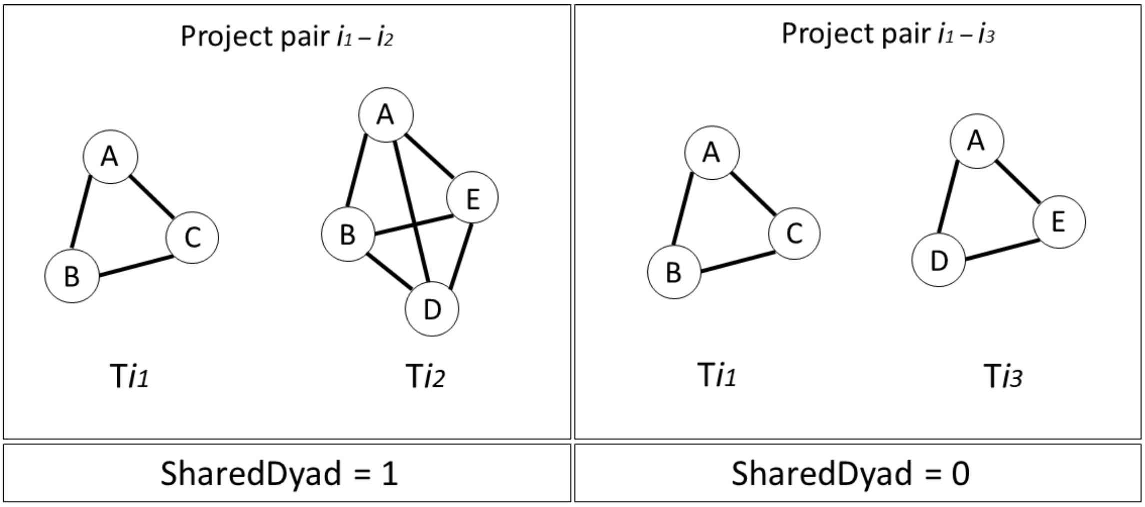

Figure 4.2. An illustrative example of the SharedDyad variable.

In the second stage of the analysis, where we concentrate exclusively on repeated collaborations with **SharedDyad** = 1, we want to test how the joint access to diverse knowledge and resources for recurring partners influences the observed degree of specialization in their repeated engagements.

Our primary explanatory variable is thus the social capital of both partners, accrued in the time between the first and the second collaboration and reflected in the measure **UniqueAlters. UniqueAlters** is a continuous variable which for each pair of organizations A-B counts how many *new unique* entities did A and B connect to since their first engagement, excluding any of the partners in the projects where A and B jointly participate. Note that we only consider first-order direct connections. While indirect links may also benefit the recipient's knowledge production, it is direct relationships that collect and process the indirect information and deliver it to the focal node (Ahuja, 2000b). The assumption here is that the total number of unique pipelines A and B can draw upon for external knowledge will influence the extent to which the pair may explore new topics when re-engaging again.

We also introduce several controls. First, we construct a dummy variable **ExtraPartners,** which takes the value of 1 when at least one of the repeated collaborations involved additional team members. This means that unless both projects i_1 and i_2 were carried out exclusively by the same pair of organizations, **ExtraPartners** will be equal to 1. Assuming that extra partners can bring in a unique set of knowledge to the collaboration, their presence in the consortium can reasonably influence the degree of specialization in repeated engagements.

Figure 4.3 illustrates the operationalization of **UniqueAlters** and **ExtraPartners,** using concrete examples. In the case of **UniqueAlters**, we can see that at the time of the second collaboration between A and B (at t=1), the pair is connected to 2 new unique organizations [D, E], to whom neither A nor B had a connection at t=0. Therefore, in the example provided **UniqueAlters** is equal to 2. In the case of **ExtraPartners** in Figure 3**,** since one of the collaborations between A and B involves an additional partner C, the dummy variable takes the value of 1.

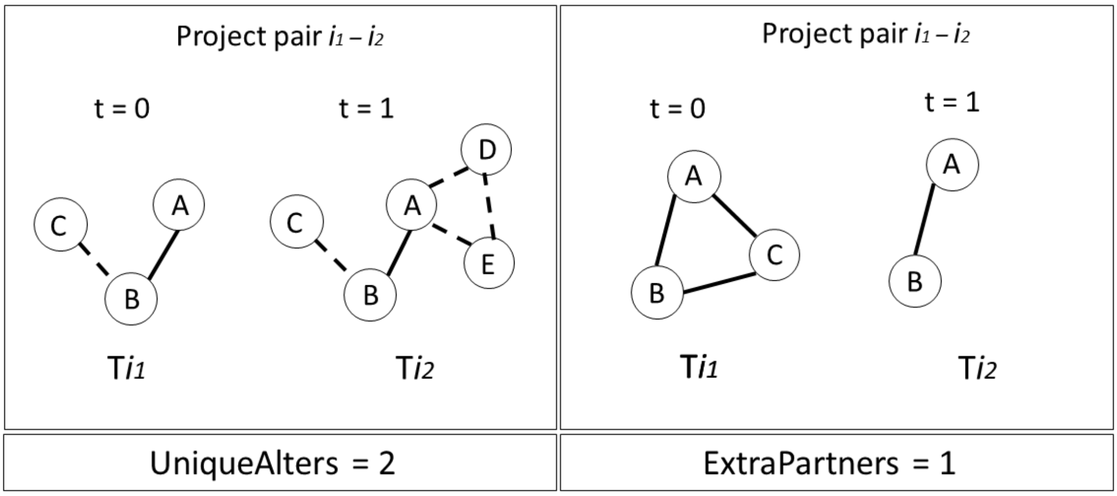

Figure 4.3. An illustrative example of the UniqueAlters and ExtraPartners variables.

We further control for the institutional characteristics of the two partners in the shared dyad. If the pair involves one public research organization (PRO) (technological institute, university-affiliated or independent research center) and a firm, **PRO-Firm** takes the value of 1, and 0 otherwise. If both members of the shared dyad are PROs, the dummy variable **PRO-PRO** takes the value of 1, and 0 otherwise. This allows us to distinguish between the behavior of a firm re-engaging with a PRO, as opposed to two PROs collaborating again. Finally, we also control for the time lag between the first and second collaboration. Since the public calls are launched on an annual basis, the variable **TimeLag** is a simple count of the number of years passed between the first and second engagement. The model to be estimated is given by:

Abstract Similarity Score = UniqueAlters + ExtraPartners + PRO-Firm + PRO-PRO + TimeLag

4.5. Results and discussion

4.5.1. Repeated engagement and specialization: descriptive results

In this section we present the results of the two-stage analysis. We begin by exploring the distribution of Abstract Similarity Score and SharedDyad (Table 4.2). One immediate observation is that the majority of abstract pairs show no significant overlap in textual content. The distribution is highly skewed, with only a small fraction of pairs being very closely related. This is not surprising since the open calls we considered target R&D collaborations from a range of sectors, and are very thematically diverse. As for team pairs, we can see that a small fraction of project pairs contains a recurrent dyad of partners.

Table 4.2. Descriptive statistics for variables Abstract Similarity Score and SharedDyad.

Statistic	**N**	**Mean**	**St.Dev.**	**Min**	**Max**
Abstract Similarity Score	18 721	0.032	0.032	0.000	0.689
SharedDyad	18 721	0.004	0.063	0.000	1.000

The two variables are positively correlated (Pearson correlation coefficient of 0.25, significant at the 1%). We also explore the distribution of the Abstract Similarity Score for pairs of projects with no dyad vs. those with at least 1 dyad in common. Figure 4.4 shows the resulting plot.

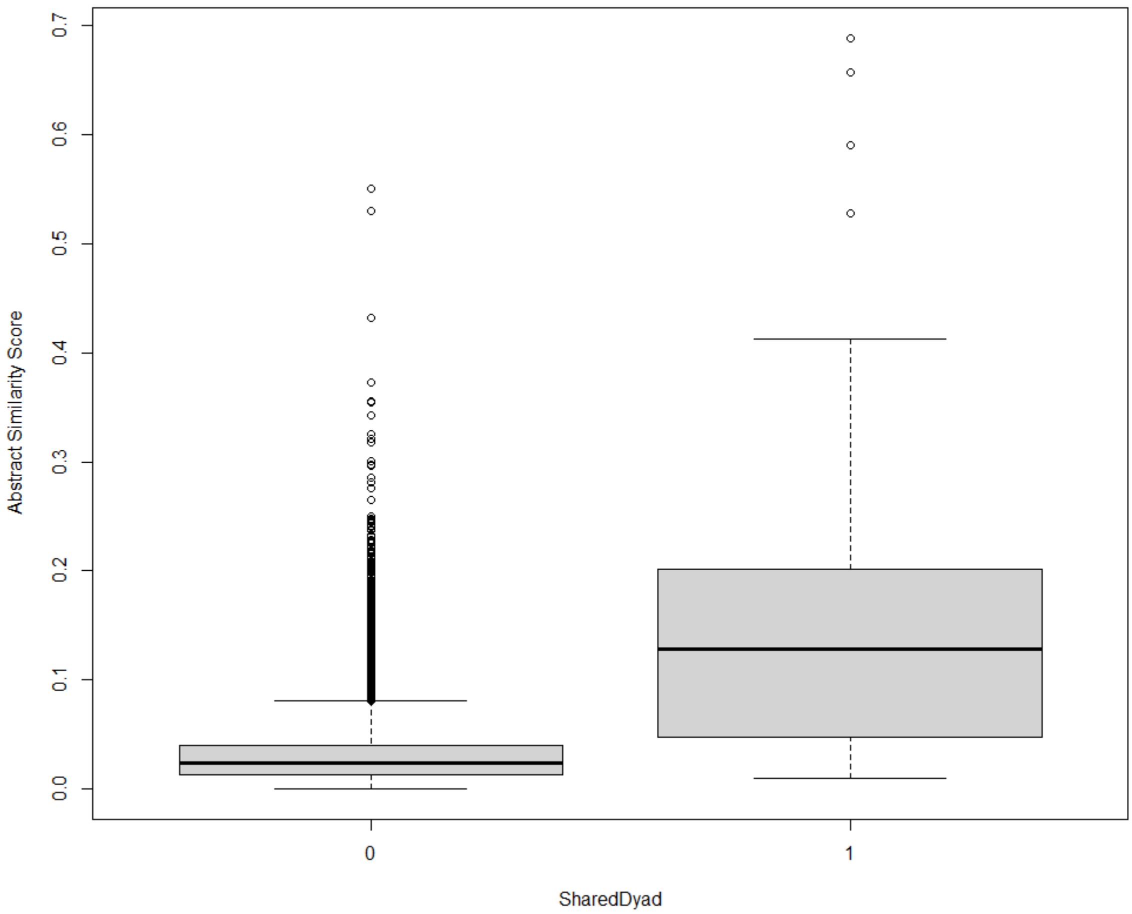

Figure 4.4. A boxplot, comparing instances of repeated collaborations (SharedDyad = 1) to the rest of project pairs (SharedDyad = 0).

What is visible from Figure 4.4 is that, on the whole, project pairs which have at least one repeating pair of partners (a shared dyad), show higher median scores of thematic similarity than projects which share only one or no partners at all. However, this generally positive relationship between repeated engagement and thematic specialization is far from straight-forward. In fact, we observe a great degree of variation across project pairs with a shared dyad. The interquartile range, which accounts for the middle 50% of scores, goes between 0.05 and 0.2, while the maximum abstract similarity score (excluding outliers) is as high as 0.4. This suggests that other factors may be at play. It appears that in some instances, collaborating pairs use subsequent R&D partnerships to extend prior work along the same topic, that is to specialize, while others do not. This provides original support to the argument put forward in the theoretical section, namely that inter-organizational network links are far from homogeneous and that only some, but not all, instances of repeated coupling between actors are associated with specialization. This implies that "getting caught up" in one type of activity after several collaborations may not necessarily be a product of the structural setting alone and the existence of strong coupling, as much as it is a product of organizations' strategic choices about how they use their strong ties.

At the same time, the bar on the left-hand side contains multiple outliers: pairs of projects which exhibit relatively high similarity, but were carried out by completely different teams.

This may be attributed to unobserved geographical proximity (i.e. organizations belong to the same cluster and therefore work on similar topics) or the presence of a common third-party, which links two distinct consortia (triadic closure) (Ter Wal, 2014). Examining social and geographical distance jointly can provide interesting insights (Singh, 2005), but falls beyond the remit of this study. Figure 4.5 shows an aspatial map of the R&D network and highlights the structural location of all repeated ties.

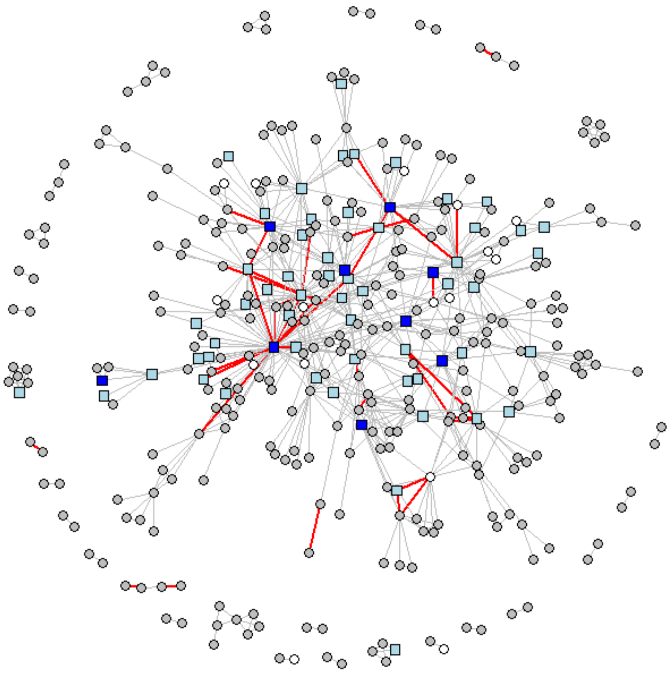

Figure 4.5. Aspatial map of the R&D network (2016-2022), highlighting repeated ties (marked in red). Nodes legend: grey circle (firm), dark blue square (technological institute), light blue square (independent or university-affiliated research center), white circle (other).

The resulting network appears centralized around the main technological institutes and several other PROs. This is consistent with studies of regional and extra-regional networks, where PROs were found to serve as intermediaries, and thus appear as frequent partners in regional collaborations (Expósito-Langa & Molina-Morales, 2010; Roediger-Schluga & Barber, 2006).

Although instances of repeated ties are relatively scarce, they appear both in the core and in the periphery. Given the positive effect of strong bonds on inter-organizational trust, their "balanced" distribution is beneficial for the flow of tacit complex knowledge across the network architecture. Figure 4.5 also showcases the institutional heterogeneity of actors involved in recurrent collaborations, which further motivates the second part of our analysis, where we explore how a dyad's access to diverse knowledge may moderate the displayed level of specialization in repeated ties.

4.5.2. The role of partners' connections for diversification

In this section, we concentrate exclusively on pairs of projects which have at least one repeated dyad (in total 75 pairs). We first display descriptive statistics (Table 4.3) and a correlation matrix (Table 4.4) of relevant variables:

Table 4.3. Descriptive statistics of the variables used in the regression analysis.

Variable	N	Mean	St. Dev.	Min	Max
Abstract Similarity Score	75	0.161	0.152	0.009	0.689
UniqueAlters	75	11.853	15.875	0	68
ExtraPartners	75	0.720	0.452	0	1
PRO-Firm	75	0.267	0.445	0	1
PRO-PRO	75	0.320	0.470	0	1
TimeLag	75	1.560	1.255	0	6

Table 4.4 Correlation matrix of the variables used in the regression analysis.

Parameter	**UniqueAlters**	**ExtraPartners**	**PRO-Firm**	**PRO-PRO**	**TimeLag**
Abstract Similarity Score	-0.40**	-0.47***	-0.24	-0.33*	0.10
UniqueAlters		0.41**	0.24	0.34*	0.20
ExtraPartners			0.38**	0.43**	-0.03
PRO-Firm				-0.41**	-0.004
PRO-PRO					-0.22

The high mean value of ExtraPartners implies that for most project pairs, the two repeating partners were not the only members in the consortium. The values for UniqueAlters vary between 0 and 68. This suggests that in some cases, the repeating partners built an extensive network of direct relationships after the first collaboration and by the time of executing the second one, the dyad had collectively accumulated 68 new unique alters in their ego network. The average time lag between the two collaborations is 1.6 years.

Examining the matrix of correlations, we note that most project pairs, where the repeated dyad is embedded in a rich network of alters, also include extra partners in the consortium. Connectivity of the repeated dyad seems to correlate with the institutional characteristics of the organizations. This can be expected since PROs tend to have a disproportionately high degree centrality. They establish numerous links with other nodes, and have sufficient human, administrative and financial capacity to maintain them (Kauffeld-Monz & Fritsch, 2013).

Table 4.5 shows the results of the beta regression. We first run a base Model 0, where we include only controls, followed by Model 1 including only the primary explanatory variable and a third model where all relevant variables are featured.

Table 4.5. Results of the beta regression. Robust standard errors appear in brackets[a].

	Dependent variable: Abstract Similarity Score		
	(Model 0)	(Model 1)	(Model 2)
UniqueAlters		-0.023***	-0.016***
		(0.007)	(0.008)
ExtraPartners	0.002		-0.048
	(0.290)		(0.286)
PRO-Firm	-0.865***		-0.513
	(0.312)		(0.335)
PRO-PRO	-0.847***		-0.487
	(0.310)		(0.348)
TimeLag	0.051		0.111
	(0.075)		(0.078)
Constant	-1.257***	-1.378***	-1.338***
	(0.206)	(0.124)	(0.205)
Observations	75	75	75
R2	0.251	0.204	0.321
Log Likelihood	70.861	67.479	72.894

[a] *Note: *p<0.1; **p<0.05; ***p<0.01*

Both control variables, reflecting the presence of one (PRO-Firm) or two (PRO-PRO) public research organizations in the dyad, have a negative coefficient, which is significant in Model 0 but not in Model 2, suggesting that PROs are generally associated with topic diversification in repeated ties. Similarly, the coefficient for UniqueAlters is negative and remains significant when controlling for the institutional heterogeneity of organizations (Model 2). This means that organizations which built an extensive network of connections and had increased access to new knowledge and ideas are also more likely to explore a different topic when re-engaging with a previous partner. Our additional control variable ExtraPartners and TimeLag do not seem to exert a significant influence on the level of specialization in repeated ties.

The results of the regression analysis suggest that organizations which are embedded in an extensive network of partners and can draw on their linkages for knowledge and resources, tend to diversify the topic of their repeated engagements. In other words, when organizations with high number of *unique* alters build strong ties, these ties exhibit more topic diversity, than strong links between isolated nodes with little connection to the rest of the network. Because of the relatively high correlation between UniqueAlters and PRO-PRO, and the fact that most central actors tend to be PROs, we cannot unequivocally attribute the “diversifying” effect to one factor alone.

On a theoretical level, the results showcase that social capital embedded in a particular linkage cannot be treated as a static asset. If one or both of the participating organizations in the dyad has a rich network of external contacts and is capable of renewing its knowledge base over time, the value of the established contact may persist longer and cycles of specialization can be followed by diversification of R&D topics. In other words, the value of strong ties may not necessarily "wear off" in an inverted U-shape the way conventional theory suggests. Moreover, these findings have implications for the framing of the proximity paradox, which seems to consider dyadic relations in isolation of the surrounding environment. When two organizations build strong ties, they do not automatically detach themselves from third parties. The conceptualization of the proximity paradox can therefore benefit from adopting a triadic approach. This will allow researchers to better understand the depreciating value of strong bonds over time. Of course, further research is needed to analyze when exactly third-party links enrich the knowledge base of a particular organization and how this influences the value of a node's persistent ties.

4.6. Conclusion

This study aims to investigate empirically the relationship between repeated inter-organizational collaboration and thematic specialization in the context of Valencia's policy-induced R&D network, and how this relationship plays out for partners with different levels of connectivity. Thus, it responds to recent calls for greater focus on networks' relational aspects and interaction processes (Madhavan & Prescott, 2017; Phelps et al., 2012), by examining specifically how organizations approach repeated collaborations. The chapter delivers several important insights.

First, it demonstrates that recurring partnerships between organizations in an R&D network are not always associated with the exploitation of the same topic, in what we call thematic specialization. Building strong bonds may also involve the exploration of new topics and the mobilization of new knowledge domains. Nevertheless, in the case of Valencia's R&D network, the latter scenario appears more likely when the partners involved are connected to a larger network of diverse contacts and can access novel knowledge and ideas. Hence, this study offers original evidence on the heterogeneity of network ties and the importance of considering the function of strong bonds between partners, given that organizations´ approach to repeated collaborations can be evidently distinct. Empirically, this chapter introduces a novel approach to measuring thematic specialization in R&D collaborations, which is based on lexical similarity of project abstracts. With regards to policymaking, the analysis is also highly relevant. When efforts are directed toward accumulating competitive advantage in prioritized areas, building strong ties between firms should be stimulated. Conversely, if the network is experiencing stagnation, exploring new themes and mobilizing novel knowledge would be far more critical. A scenario like this calls for investment in partnerships that enroll a broader range of organizations (including PROs) with a rich network of contacts, both local and extra-regional, in order to diversify the thematic focus of R&D collaborations, and avoid further specialization in declining industries.

Finally, this study is not without limitations. The most significant one concerns the operationalization of our dependent variable, which builds on the lexical similarity of project abstracts. Since it is plausible that two abstracts describe the same area of research through different terminology, it would be beneficial to repeat the analysis using an advanced semantic similarity method, which is capable of interpreting the meaning of textual information. In addition, when constructing our primary independent variable for network connectivity, we consider only links to other unique nodes within the same network. In reality, actors may be able to access external knowledge through complementary linkages in parallel unobserved formal or informal networks. Both of these limitations offer promising avenues for further research. More importantly, we believe that doubling down on efforts to examine the nature and content of inter-organizational ties can be particularly beneficial for fleshing out the big questions surrounding the co-evolution of network structures and knowledge flow.

Chapter 5. General conclusion

The objective of this thesis consisted in explaining the dynamics of inter-organizational knowledge networks, and more specifically the antecedents and implications of strong ties.

This chapter provides a concise overview of the main theoretical, methodological and policy contributions stemming from the research conducted. It explains the principal limitations of the study and suggests several avenues for future research.

5.1. Research contributions

5.1.1. Theoretical

The results of this thesis enrich two separate streams of literature: the one on knowledge network dynamics, and the one on strategic management.

On the one hand, we show in chapter 3 that the evolution of knowledge networks is driven by different forms of social proximity. In the existing scholarly discourse, social proximity is operationalized by prior successful collaborations, whereas we demonstrate empirically that an “unsuccessful relationship” between two partners, in the form of failed project application, is also a form of relational embeddedness which can stimulate re-engagement and influence the structural evolution of the network. Furthermore, we illustrate that this form of relational inertia between unsuccessful partners is stronger when they are cognitively distant. To our knowledge, this relationship has not been tested empirically in prior studies.

In chapter 4 we advance the discussion further by showing that structurally equivalent network ties can assume fundamentally distinct roles, by leading either to thematic specialization or diversification. Thus, we directly contribute to the debate on the benefits and drawbacks of strong ties. According to the proximity paradox, the trust and familiarity which cause one-time partners to re-engage and share knowledge, are the same factors that cultivate a sort of intellectual comfort zone in the long-run and block the entry of new ideas and perspectives (Boschma & Frenken, 2010; Broekel & Boschma, 2012). Thus, building strong ties is assumed to be beneficial in the short-run but progressively risky in the long-run as partners become too embedded in existing relationships.

The results of this doctoral work suggest that the value of strong ties may not necessarily “wear off” the way conventional theory suggests in an inverted U-shape (Gargiulo & Benassi, 2000; Masciarelli et al., 2010) . By demonstrating that some actors use strong ties to specialize in a single domain, while others to diversify into distinct topics, we prove that the value of strong ties will depend at least partially on the way actors choose to leverage their strong connections. In other words, some organizations' entrapment in a single activity may not be solely attributed to the inherent structural framework and the existence of strong ties. Instead, the outcome likely hinges on the strategic decisions these organizations make regarding the exploitation of their repeated collaborations.

Moreover, these findings have implications for the framing of the proximity paradox, which seems to consider dyadic relations in isolation of the surrounding environment. When two organizations build strong ties, they do not by default detach themselves from third parties. The conceptualization of the proximity paradox can therefore benefit from adopting a triadic approach. This will allow researchers to better understand the depreciating value of strong bonds over time, that is: when exactly repetitive engagements start to “trap” R&D partners into sub-optimal innovative performance, and how the contact with third-party organizations may dampen this negative effect.

With regards to the strategic management literature, our contributions are two-fold. First, in chapter 3 we show that collaborating organizations often engage in cost optimization strategies. Rather than abandoning unsuccessful partnerships, organizations appear more willing to try and benefit from the initial investment of creating a partnership in the first place. Finding appropriate partners and establishing a working connection with them can be very time-consuming, so organizations look for ways to recover those transaction costs by re-engaging with familiar partners, even if their initial cooperation was unsuccessful in obtaining funding. While organizations’ tendency to re-apply following a rejection in competitive R&D schemes was studied previously (Barajas & Huergo, 2010), our examination adds a new layer of insight to the existing body of knowledge: re-applications often happens with the same partners. This means organizations seek to recover a combination of costs: that of developing a competitive project proposal, *and* the search cost of finding suitable partners.

Second, in chapter 4 we further show that organizations re-engage with prior partners for different purposes. Some choose to exploit the same topic as in their first collaboration, thus deepening their expertise in a single domain, while others embark on the exploration of fundamentally distinct research topics. To the best of our knowledge, this heterogeneity in organizations’ collaborative behavior has not been demonstrated empirically before. Moreover, we reveal that organizations which build an extensive network of connections and have increased access to new knowledge and ideas, especially universities and research centers, are more likely to explore a different topic when re-engaging with a previous partner.

5.1.2. Methodological

From a methodological standpoint, the doctoral work presents several noteworthy contributions. First, it employs a unique dataset, which was curated exclusively for the purposes of this research, and therefore it has not been used before. The manual data collection increased the level of detail in the final sample we use, featuring information which is often absent in conventional repositories, such as records on rejected R&D projects, and a succinct textual description of most approved projects. The regional component is equally valuable. Systematic information on sub-national subsidized R&D collaborations is often difficult to find and if available – is typically lacking in consistency (exceptions include Broekel & Mueller (2018), Tsouri & Pegoretti (2020)). The longitudinal nature of the data is another important asset. Early studies on network dynamics tend to view networks from a static perspective, capturing inter-relations as a snapshot in time. Although in recent years longitudinal studies have received greater attention, there is still much to learn from the dynamic exploration of networks’ temporal evolution (Ter Wal & Boschma, 2009b).

Second, the thesis presents a novel operationalization of a familiar concept – social proximity. In chapter 3, we build three different variables which reflect the possible scenarios of prior interaction between organizations – failed partnership, successful partnership, or a mixture of both. To our knowledge, this differentiation has not been applied previously in the context of network dynamics studies, and is particularly valuable because it highlights the idea that social proximity does not exclusively arise from successful collaborations, but can also be cultivated through experiences of failure.

Finally, the use of natural language processing techniques to compare project descriptions and build a measure of thematic similarity constitutes another contribution. In chapter 4 we employ cosine similarity to assess the degree of lexical overlap in pairs of project abstracts. So far, machine learning techniques have been mobilized primarily to assess patent novelty and the departure from established ideas, generally denoted as "breakthroughs". Our intention, however, is to harness the potential of NLP methods in a different setting: to examine the tendency of organizations to stick to the same R&D topic when collaborating repeatedly with the same partner, as opposed to pivoting away in new research directions. In that sense, we are not measuring novelty or originality, but focusing more on the organizational behavior of partners in recurring collaborations. Our goal is to shed light on organizations' deliberate choices and their propensity to thematically specialize when re-engaging with previous partners. This constitutes a novel deviation from the above-mentioned examples of NLP use.

5.1.3. Policy

The findings of this doctoral study also hold significant implications for practitioners working in various levels of governance. Given that the empirical analysis is centered on R&D partnerships, which were directly shaped by government-designed instruments under the smart specialization framework, we are well-positioned to provide a range of specific and more general insights into the workings of policy-backed R&D networks.

Context-specific recommendations

We begin by providing targeted context-specific recommendations, which concern the particular case of Valencia's inter-organizational network.

Given the high spatial concentration of R&D partnerships along the three key cities – Valencia, Castellon and Alicante, regional policymakers may consider introducing *soft* measures that stimulate urban-rural engagement. Incentivizing the participation of new partners from remote municipalities, which are currently absent from the map of interactions, can help promote regional equity and avoid deepening intra-regional divides. Since we consider the entire list of project applications (both approved and rejected), we can confidently conclude that the problem lies not in a biased allocation of funding, but in the lack of incentives for rural partners to pitch a proposal in competitive R&D programs. Further research is needed to determine whether the lack of incentives is coupled with information asymmetries or other additional factors.

Second, the central network position occupied by public research organizations in Valencia and their consistent appearance as partners in instances of repeated collaborations leaves little space for firms to influence the knowledge dynamics of the regional network. Perhaps, more should be done to remove potential barriers to re-application which appear to discourage firms more than universities or research centers. Of course, we look at a relatively narrow 7-year period, and perhaps firms do not have sufficient resources to reapply immediately following a rejection. Nevertheless, this underscores the importance of policymakers taking proactive measures to intervene and create a more accessible and equitable landscape for all participants.

Broader policy implications

The implications of this study extend well beyond the Valencian case, given that funding competitive trans-sectoral R&D schemes has become a signature policy instrument for many regional and national governments. The analysis in chapter 3 highlighted the important role these instruments play in generating a “hidden” network of partners, whose joint experience, even when negative (i.e. project rejection), serves as a strong motivation for further re-engagement.

The observed propensity toward cost optimization implies that rejected partners are likely to pursue alternative ways of funding their project plans. This could involve re-applying to the same program, if possible, or to a different national or even international R&D scheme. In that sense, as we argued earlier, the social proximity generated by government-run programs cannot be seen as an “expendable” asset that disappears after a project application has been denied funding. Rather, it persists as an unobserved factor driving the evolution of interorganizational relations over time.

Since the policy instruments we examined form part of the Valencian smart specialization strategy, it is fitting to highlight potential implications for the policy itself. In Chapter 4 we showed that organizations’ approach to re-engagement can be evidently distinct, but the inclusion of PROs can increase the chance of topic diversification. This suggests that policymakers may be better advised to stimulate public-private partnerships when new avenues for R&D are needed to break out of unproductive cycles. This is especially crucial when the thematic research priorities outlined in the strategy differ from what has been traditionally emphasized in the region. On the other hand, our analysis demonstrated that recurring partnerships between firms are more bound to stick to the exploitation of the same topic. This phenomenon could offer its own set of advantages, including greater *thematic* specialization, which is desirable if it goes along the *regional* specialization priorities established in the strategy. Thus, depending on the status of the regional network, policymakers will need to adjust their interventions appropriately, to either stimulate or deter further thematic specialization in extant R&D topics.

5.2. Limitations

As with any research work, this thesis is not without limitations. We address the most important ones in the section below:

5.2.1. Scope and generalizability considerations

A primary limitation of our study is its exclusive focus on a single network. While the narrow scope allowed us to gather greater level of detail, which is otherwise lacking in regional R&D datasets, our work refers to Valencia region only and considers a limited set of policy instruments in a restricted time period. The findings can therefore be influenced by Valencia's unique context and idiosyncrasies. Nevertheless, we showcase that the region we study is comparable to many others in south and east Europe, which are characterized by fragmented business fabric, high prevalence of micro and small enterprises, dominance of more traditional sectors and an active core of public research organizations. Our conclusions are therefore relevant beyond the Valencian case, but should not be used to overly generalize the broad phenomenon of inter-organizational network formation and evolution.

5.2.2. Unobserved latent ties

One of the most common limitations in any type of network study is the presence of latent ties. These are unobserved or hidden connections between nodes, which are not directly visible or documented in the data being analyzed. When defining the scope of our analysis, we chose to concentrate on research ties which emerged as a result of three specific regional policy instruments aimed at stimulating collaborative R&D (AVI's programs "Strategic projects in cooperation" and "Consolidation of the business value chain", and IVACE's program "R&D in cooperation"). In reality, the complexity of partners' relations is far greater. Thus, for instance, it is possible that two organizations which appear as first-time collaborators in the network of interest have had multiple prior engagements in unobserved networks and are in fact quite familiar with each other. This complicates the study of social proximity, since latent ties can also generate trust between partners. This is why, we must be cautious not to overemphasize the absence of a specific tie as an indicator of the absence of any type of relationship. We stress that the conclusions derived in chapters 4 and 5 must be interpreted in the context of the R&D network we are analyzing, keeping in mind that organizations may be connected through different types of formal and informal relationships we cannot fully account for.

5.2.3. External factors

Last but not least, the timeframe of our analysis spanning from 2016 to 2022, includes a particularly turbulent period, marked by a global pandemic, which fueled deep economic uncertainty across virtually all sectors and industries. It is therefore plausible that outside factors affected organizations' collaborative choices and their strategic approach to R&D collaborations. These external circumstances were not directly accounted for in our analysis but should be studied separately, since network behavior is likely to differ during periods of crisis or uncertainty (Tsouri, 2019).

5.3. Future lines of research

The conclusions presented in this chapter open new possibilities for future research.

On the one hand, we showcase in chapter 3 that rejected project applications serve as a strong predictor of future tie formation, but we do not know why exactly organizations choose to re-engage with the same partner following a negative experience. Much of the nuances in organizational behavior can only be understood by incorporating qualitative methods, such as surveys or interviews. Investigating the learning processes that take place after a failed project application is similarly important, since organizations may acquire valuable adaptive skills and knowledge, which positively influence the outcomes of their future collaborations.

Furthermore, we noted the high presence of micro and small enterprises in the regional network. These types of local actors often face challenges and constraints in competing in national or international excellence programs. Thus, at least in theory, the accessibility of regional funding makes it possible for them to develop collaborative R&D ties. Yet, in this thesis we do not directly test whether bigger well-established firms engage with smaller SMEs, and whether the network membership as such contributes to increasing the knowledge stock of the smaller actors and their overall competitiveness. In essence, the sheer diversity of organizations, which characterizes our network study, does not inherently guarantee that these actors will connect with one another and tap into the knowledge opportunities the network offers.

On the other hand, in chapter 4 we establish a relationship between strong ties and thematic specialization. However, we do not directly assess how the consistent exploitation of the same R&D topic influences the final outcome of the collaboration, in comparison with partners who venture into new lines of work (diversification). This next step of analysis can shed light on the merits of strong ties, highlighting just when and how the social proximity between partners benefits their innovative knowledge production.

Moreover, we concentrate on the topic and content of project abstracts, but the advancement of text analysis methods may soon make it easier for researchers to gather even more elaborate information on the specific resources mobilized in a given collaboration, including knowledge, material, financial, or social assets (Revet, 2022; Shibayama et al., 2012). This will highlight not only organizations' topic selection, but potentially illuminate the underlying process of knowledge exchange between partners. Network resource flows have long been a black box in the study of inter-personal and inter-organizational relations, partly due to data constraints and methodological challenges. The use of NLP may help solve the latter of the two limitations, bringing us closer to a comprehensive understanding of networks' intricate internal processes.

In conclusion, this doctoral work builds on the pillars of classical network theory but makes use of unique data and innovative methodological approaches to flesh out the details of how inter-organizational network structure and knowledge exchange processes coevolve. We hope that the contributions of our work, encapsulated in the pages of this thesis, will prove meaningful and valuable to scholars and practitioners alike, and that they will inspire future researchers to dive deeper into this fascinating and ever-evolving discipline of science.

Reference list

Ahuja, G. (2000a). The duality of collaboration: inducements and opportunities in the formation of interfirm linkages. *Strategic Management Journal*, *21*(3), 317–343. https://doi.org/https://doi.org/10.1002/(SICI)1097-0266(200003)21:3<317::AID-SMJ90>3.0.CO;2-B

Ahuja, G. (2000b). Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. *Administrative Science Quarterly*, *45*(3), 425–455. https://doi.org/10.2307/2667105

Ahuja, G., Soda, G., & Zaheer, A. (2012). Introduction to the Special Issue: The Genesis and Dynamics of Organizational Networks. *Organization Science*, *23*(2), 434–448. http://www.jstor.org/stable/41429345

Al-Tabbaa, O., & Ankrah, S. (2016). Technological Forecasting & Social Change Social capital to facilitate ‘ engineered ’ university – industry collaboration for technology transfer : A dynamic perspective. *Technological Forecasting & Social Change*, *104*, 1–15. https://doi.org/10.1016/j.techfore.2015.11.027

Amoroso, S. (2014). The hidden costs of R&D collaboration. *IPTS Working Papers on Corporate R&D and Innovation*, *02*.

Ankrah, S. N., Burgess, T. F., Grimshaw, P., & Shaw, N. E. (2013). Asking both university and industry actors about their engagement in knowledge transfer: What single-group studies of motives omit. *Technovation*, *33*(2), 50–65. https://doi.org/https://doi.org/10.1016/j.technovation.2012.11.001

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*, 495–519. https://doi.org/https://doi.org/10.1111/j.1435-5957.2007.00132.x

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

Balland, P.-A., Belso-Martínez, J. A., & Morrison, A. (2016). The dynamics of technical and business knowledge networks in industrial clusters: Embeddedness, status, or proximity? *Economic Geography*, *92*(1), 35–60. https://doi.org/10.1080/00130095.2015.1094370

Balland, P.-A., Boschma, R., & Frenken, K. (2022). Proximity, Innovation and Networks: A Concise Review and Some Next Steps. In A. Torre & D. Gallaus (Eds.), *Handbook of Proximity Relations* (pp. 70–81). Edward Elgar Publishing.

Balland, P.-A., Boschma, R., & Ravet, J. (2019). Network dynamics in collaborative research in the EU, 2003–2017. *European Planning Studies*, *27*(9), 1811–1837. https://doi.org/10.1080/09654313.2019.1641187

Balland, P.-A., De Vaan, M., & Boschma, R. (2013). The dynamics of interfirm networks along the industry life cycle. *Journal of Economic Geography*, *13*(5), 741–765. https://www.jstor.org/stable/26158687

Balsmeier, B., Assaf, M., Chesebro, T., Fierro, G., Johnson, K., Johnson, S., Li, G. C., Lück, S., O'Reagan, D., Yeh, B., Zang, G., & Fleming, L. (2018). Machine learning and natural language processing on the patent corpus: Data, tools, and new measures. *Journal of Economics and Management Strategy*, *27*(3), 535–553. https://doi.org/10.1111/jems.12259

Barabási, A.-L. (2009). Scale-Free Networks: A Decade and Beyond. *Science*, *325*(5939), 412–413. https://doi.org/10.1126/science.1173299

Barabási, A.-L., & Bonabeau, E. (2003). Scale-Free Networks. *Scientific American*, *288*(5), 60–69. http://www.jstor.org/stable/26060284

Barajas, A., & Huergo, E. (2010). International R&D cooperation within the EU Framework Programme: empirical evidence for Spanish firms. *Economics of Innovation and New Technology*, *19*(1), 87–111. https://doi.org/10.1080/10438590903016492

Barnes, T., Pashby, I., & Gibbons, A. (2002). *Industry interaction : A multi-case evaluation of collaborative R&D projects*. *20*(3), 272–285.

Baum, J. A. C., Calabrese, T., & Silverman, B. S. (2000). Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology. *Strategic Management Journal*, *21*(3), 267–294. https://doi.org/https://doi.org/10.1002/(SICI)1097-0266(200003)21:3<267::AID-SMJ89>3.0.CO;2-8

Becerra, M., Lunnan, R., & Huemer, L. (2008). Trustworthiness, risk, and the transfer of tacit and explicit knowledge between alliance partners. *Journal of Management Studies*, *45*(5), 1024. https://doi.org/10.1111/j.1467-6486.2008.00788.x

Borgatti, S. P. (2009). *Social Network Analysis, Two-Mode Concepts in BT - Encyclopedia of Complexity and Systems Science* (R. A. Meyers (ed.); pp. 8279–8291). Springer New York. https://doi.org/10.1007/978-0-387-30440-3_491

Borgatti, S. P., & Foster, P. C. (2003). The Network Paradigm in Organizational Research: A Review and Typology. *Journal of Management*, *29*(6), 991–1013. https://doi.org/10.1016/S0149-2063_03_00087-4

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61–74. https://doi.org/10.1080/0034340052000320887

Boschma, R., & Frenken, K. (2010). The spatial evolution of innovation networks: A proximity perspective. In R. Boschma & R. Martin (Eds.), *The Handbook of Evolutionary Economic Geography* (pp. 120–138). Edward Elgar Publishing. https://doi.org/10.4337/9781849806497.00012

Boschma, R., & Ter Wal, A. L. J. (2007). Knowledge networks and innovative performance in an industrial district: the case of a footwear district in the south of Italy. *Industry and Innovation*, *14*(2), 177–199. https://doi.org/10.1080/13662710701253441

Broekel, T., Balland, P.-A., Burger, M., & van Oort, F. (2014). Modeling knowledge networks in economic geography: a discussion of four methods. *The Annals of Regional Science*, *53*(2), 423–452. https://doi.org/10.1007/s00168-014-0616-2

Broekel, T., & Boschma, R. (2012). Knowledge networks in the Dutch aviation industry: The proximity paradox. *Journal of Economic Geography*, *12*(2), 409–433. https://doi.org/10.1093/jeg/lbr010

Broekel, T., & Hartog, M. (2013). Determinants of cross-regional R&D collaboration networks : an application of exponential random graph models. In T. Scherngell (Ed.), *Advances in Spatial Science. The geography of networks and R&D collaborations* (pp. 49–70).

Broekel, T., & Mueller, W. (2018). Critical links in knowledge networks – What about proximities and gatekeeper organisations? *Industry and Innovation*, *25*(10), 919–939. https://doi.org/10.1080/13662716.2017.1343130

Bruneel, J., D'Este, P., & Salter, A. (2010). Investigating the factors that diminish the barriers to university-industry collaboration. *Research Policy*, *39*(7), 858–868. https://doi.org/10.1016/j.respol.2010.03.006

Bstieler, L., Hemmert, M., & Barczak, G. (2017). The changing bases of mutual trust formation in inter-organizational relationships: A dyadic study of university-industry research collaborations. *Journal of Business Research*, *74*, 47–54. https://doi.org/https://doi.org/10.1016/j.jbusres.2017.01.006

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349–399. https://doi.org/10.1086/421787

Caloffi, A., Rossi, F., & Russo, M. (2015). What Makes SMEs more Likely to Collaborate? Analysing the Role of Regional Innovation Policy. *European Planning Studies*, *23*(7), 1245–1264. https://doi.org/10.1080/09654313.2014.919250

Camagni, R. P. (1995). The concept of innovative milieu and its relevance for public policies in European lagging regions. *Papers in Regional Science*, *74*(4), 317–340. https://doi.org/https://doi.org/10.1111/j.1435-5597.1995.tb00644.x

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

Cantner, U., Hinzmann, S., & Wolf, T. (2017). *The Coevolution of Innovative Ties, Proximity, and Competencies: Toward a Dynamic Approach to Innovation Cooperation*. https://doi.org/10.1007/978-3-319-45023-0_16

Cantner, U., & Meder, A. (2007). Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination*, *2*(1), 45–65. https://doi.org/10.1007/s11403-007-0018-y

Chandrasekaran, D., & Mago, V. (2021). Evolution of semantic similarity - A survey. *ACM Computing Surveys*, *54*(2), 1–37.

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

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, *94*, S95–S120. http://www.jstor.org/stable/2780243

Cooke, P. (2003). Regional Innovation and Learning Systems, Clusters, and Local and Global Value Chains. In J. Bröcker, D. Dohse, & R. Soltwedel (Eds.), *Innovation Clusters and Interregional Competition. Advances in Spatial Science* (pp. 28–51). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-24760-9_3

Cooke, P. (2008). Regional innovation systems: origin of the species. *International Journal of Technological Learning, Innovation and Development*, *1*(3), 393–409. https://doi.org/10.1504/IJTLID.2008.019980

Cunningham, P., Edler, J., Flanagan, K., & Larédo, P. (2013). Innovation policy mix and instrument interaction: A review. In *NESTA Compendium of Evidence on the Effectiveness of Innovation Policy Intervention* (Issue 13/20). http://www.innovation-policy.net/compendium/

Cunningham, P., & Gök, A. (2012). The impact and effectiveness of policies to support collaboration for R&D and innovation. In *NESTA Compendium of Evidence on the Effectiveness of Innovation Policy Intervention* (Issue 12/06). http://www.innovation-policy.net/compendium/

Cunningham, P., & Ramlogan, R. (2012). *The effects of innovation network policies* (12/04; NESTA Compendium of Evidence on the Effectiveness of Innovation Policy Intervention).

D'Este, P., Guy, F., & Iammarino, S. (2013). Shaping the formation of university-industry research collaborations: What type of proximity does really matter? *Journal of Economic Geography*, *13*(4), 537–558. https://doi.org/10.1093/jeg/lbs010

Dahlander, L., & McFarland, D. A. (2013). Ties that last: Tie formation and persistence in research collaborations over time. *Administrative Science Quarterly*, *58*(1), 69–110. https://doi.org/10.1177/0001839212474272

Das, T. K., & Teng, B.-S. (2002). Alliance Constellations: A Social Exchange Perspective. *The Academy of Management Review*, *27*(3), 445–456. https://doi.org/10.2307/4134389

Diestre, L., & Rajagopalan, N. (2012). Are all "sharks" dangerous? new biotechnology ventures and partner selection in R&D alliances. *Strategic Management Journal*, *33*(10), 1115–1134. https://doi.org/10.1002/smj.1978

Doern, G. B., & Stoney, C. (2009). Federal Research and Innovation Policies and Canadian Universities: A Framework for Analysis. In G. B. Doern & C. Stoney (Eds.), *Research and Innovation Policy: Changing Federal Government - University Relations* (pp. 3–34). University of Toronto Press. https://doi.org/doi:10.3138/9781442697478

Doloreux, D., & Parto, S. (2005). Regional innovation systems: Current discourse and unresolved issues. *Technology in Society*, *27*(2), 133–153. https://doi.org/10.1016/j.techsoc.2005.01.002

Doz, Y. L. (1996). The evolution of cooperation in strategic alliances: Initial conditions or learning processes? *Strategic Management Journal*, *17*(S1), 55–83. https://doi.org/https://doi.org/10.1002/smj.4250171006

Dyer, J. H., & Singh, H. (2011). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *The Academy of Management Review*, *23*(4), 660–679. https://doi.org/https://doi.org/10.2307/259056

Enger, S. G. (2018). Closed clubs: Network centrality and participation in Horizon 2020. *Science and Public Policy*, *45*(6), 884–896. https://doi.org/10.1093/scipol/scy029

Enger, S. G., & Castellacci, F. (2016). Who gets Horizon 2020 research grants? Propensity to apply and probability to succeed in a two-step analysis. *Scientometrics*, *109*(3), 1611–1638. https://doi.org/10.1007/s11192-016-2145-5

European Commission, D.-G. for R. and I. (2021). *Horizon Europe, budget : Horizon Europe - the most ambitious EU research & innovation programme ever*. https://doi.org/https://data.europa.eu/doi/10.2777/202859

Expósito-Langa, M., & Molina-Morales, F. X. (2010). How relational dimensions affect knowledge redundancy in industrial clusters. *European Planning Studies*, *18*(12), 1975–1992. https://doi.org/10.1080/09654313.2010.515817

Feng, S. (2020). The proximity of ideas: An analysis of patent text using machine learning. *PLOS ONE*, *15*(7), e0234880. https://doi.org/10.1371/journal.pone.0234880

Ferrari, S., & Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics*, *31*(7), 799–815. https://doi.org/10.1080/0266476042000214501

Ferriani, S., Fonti, F., & Corrado, R. (2012). The social and economic bases of network multiplexity: Exploring the emergence of multiplex ties. *Strategic Organization*, *11*(1), 7–34. https://doi.org/10.1177/1476127012461576

Florida, R. (1995). Toward the Learning Region. *Futures*, *27*(5), 527–536. https://api.semanticscholar.org/CorpusID:67767875

Foray, D. (2017). Chapter 2 - The Economic Fundamentals of Smart Specialization Strategies. In S. Radosevic, A. Curaj, R. Gheorghiu, L. Andreescu, & I. B. T.-A. in the T. and P. of S. S. Wade (Eds.), *Advances in the Theory and Practice of Smart Specialization* (pp. 37–50). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-804137-6.00002-4

Foray, D., David, P. A., & Hall, B. (2009). Smart Specialisation – The Concept. In *European Commission* (Vol. 9, Issue June). “Knowledge for Growth” Expert Group.

Foray, D., Morgan, K., & Radosevic, S. (2018). *The role of smart specialization in the EU research and innovation policy landscape*. https://ec.europa.eu/regional_policy/en/information/publications/brochures/2018/the-role-of-smart-specialisation-in-the-eu-research-innovation-policy-landscape

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabási, A.-L. (2018). Science of science. *Science*, *359*(6379), eaao0185. https://doi.org/10.1126/science.aao0185

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/https://doi.org/10.1016/0378-8733(78)90021-7

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, *41*(5), 685–697. https://doi.org/10.1080/00343400601120296

Fritsch, M., & Kauffeld-Monz, M. (2009). 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*(1), 21–38. https://doi.org/10.1007/s00168-008-0245-8

Gargiulo, M., & Benassi, M. (1999). *The Dark Side of Social Capital BT - Corporate Social Capital and Liability* (R. T. A. J. Leenders & S. M. Gabbay (eds.); pp. 298–322). Springer US. https://doi.org/10.1007/978-1-4615-5027-3_17

Gargiulo, M., & Benassi, M. (2000). Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital. *Organization Science*, *11*(2), 183–196. http://www.jstor.org/stable/2640283

Generalitat-Valenciana. (2016). *Estrategia de Especialización Inteligente para la investigación e innovación en la Comunitat Valenciana*.

Generalitat-Valenciana. (2019). *Evaluación intermedia de la estrategia de especialización inteligente para la investigación e innovación en la Comunitat Valenciana 2014-2020 Volumen II*.

Georghiou, L. (2002). Impact and Additionality of Innovation Policy. *Six Countries Programme on Innovation Policy and Sustainable Development*.

Gilsing, V., Nooteboom, 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*(10), 1717–1731. https://doi.org/https://doi.org/10.1016/j.respol.2008.08.010

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

Glückler, J., & Doreian, P. (2016). Editorial: social network analysis and economic geography—positional, evolutionary and multi-level approaches. *Journal of Economic Geography*, *16*(6), 1123–1134. https://doi.org/10.1093/jeg/lbw041

Goerzen, A. (2007). Alliance networks and firm performance: the impact of repeated partnerships. *Strategic Management Journal*, *28*, 487–509. https://doi.org/10.1002/smj

Gomaa, W. H., & Fahmy, A. A. (2013). Article: A survey of text similarity approaches. *International Journal of Computer Applications*, *68*(13), 13–18.

Graf, H., & Krüger, J. J. (2011). The Performance of Gatekeepers in Innovator Networks. *Industry and Innovation*, *18*(1), 69–88. https://doi.org/10.1080/13662716.2010.528932

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Gulati, R. (1995). Does familiarity breed trust? The implications of repeated ties for contractual choice in alliance. *Academy of Management*, *38*(1), 85–112.

Gulati, Ranjay, & Gargiulo, M. (1999). Where do interorganizational networks come from? *American Journal of Sociology*, *104*(5), 1439–1493. http://www.jstor.org/stable/10.1086/210179

Hagedoorn, J. (2002). Inter-firm R&D partnerships: An overview of major trends and patterns since 1960. *Research Policy*, *31*(4), 477–492. https://doi.org/10.1016/s0048-7333(01)00120-2

Hagedoorn, J., Link, A. N., & Vonortas, N. S. (2000). Research partnerships. *Research Policy*, *29*(4), 567–586. https://doi.org/https://doi.org/10.1016/S0048-7333(99)00090-6

Hammoud, Z., & Kramer, F. (2020). Multilayer networks: aspects, implementations, and application in biomedicine. *Big Data Analytics*, *5*(1), 2. https://doi.org/10.1186/s41044-020-00046-0

Han, J., Kamber, M., & Pei, J. (2012). Getting to know your data. In J. Han, M. Kamber, & J. B. T.-D. M. (Third E. Pei (Eds.), *The Morgan Kaufmann Series in Data Management Systems* (pp. 39–82). Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/B978-0-12-381479-1.00002-2

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, *44*(1), 82–111. https://doi.org/10.2307/2667032

Heringa, P. W., Hessels, L. K., & van der Zouwen, M. (2016). The influence of proximity dimensions on international research collaboration: an analysis of European water projects. *Industry and Innovation*, *23*(8), 753–772. https://doi.org/10.1080/13662716.2016.1215240

Hernán, R., Marín, P. L., & Siotis, G. (2003). An empirical evaluation of the determinants of Research Joint Venture Formation. *The Journal of Industrial Economics*, *51*(1), 75–89. https://doi.org/https://doi.org/10.1111/1467-6451.00192

Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, *28*(4), 331–343. https://doi.org/https://doi.org/10.1002/smj.582

Howells, J. R. L. (2002). Tacit knowledge, innovation and economic geography. *Urban Studies*, *39*(5–6), 871–884. https://doi.org/10.1080/00420980220128354

Hünermund, P., Lopes Bento, C., & Pellens, M. (2022). The Effect of Publicly Co-funded Industry-science Collaboration on Scientific Production. *DRUID Working Paper Series*.

Hyvärinen, J., & Rautiainen, A.-M. (2007). Measuring additionality and systemic impacts of public research and development funding — the case of TEKES, Finland. *Research Evaluation*, *16*(3), 205–215. https://doi.org/10.3152/095820207X235115

Inkpen, A. C., & Tsang, E. W. K. (2005). Social Capital, Networks, and Knowledge Transfer. *Academy of Management Review*, *30*(1), 146–165. https://doi.org/10.5465/amr.2005.15281445

Inoue, H. (2015). Evidence for a Creative Dilemma Posed by Repeated Collaborations. *PLOS ONE*, *10*(9), e0137418. https://doi.org/10.1371/journal.pone.0137418

Inoue, H., & Liu, Y.-Y. (2015). Revealing the Intricate Effect of Collaboration on Innovation. *PLOS ONE*, *10*(3), e0121973. https://doi.org/10.1371/journal.pone.0121973

Juhász, S. (2021). Spinoffs and tie formation in cluster knowledge networks. *Small Business Economics*, *56*(4), 1385–1404. https://doi.org/10.1007/s11187-019-00235-9

Kaplan, S., & Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, *36*(10), 1435–1457. https://doi.org/https://doi.org/10.1002/smj.2294

Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, *31*(1), 31–43. https://doi.org/10.1007/BF02018100

Kauffeld-Monz, M., & Fritsch, M. (2013). Who are the knowledge brokers in regional systems of innovation? A multi-actor network analysis. *Regional Studies*, *47*(5), 669–685. https://doi.org/10.1080/00343401003713365

Kelly, B., Papanikolaou, D., Seru, A., & Taddy, M. (2021). Measuring technological innovation over the long run. *American Economic Review: Insights*, *3*(3), 303–320. https://doi.org/10.1257/aeri.20190499

Lavie, D. (2006). The Competitive Advantage of Interconnected Firms: An Extension of the Resource-Based View. *The Academy of Management Review*, *31*(3), 638–658. http://www.jstor.org/stable/20159233

Lundvall, B.-Å. (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. Pinter Publishers.

Luukkonen, T. (2000). Additionality of EU framework programmes. *Research Policy*, *29*(6), 711–724. https://doi.org/https://doi.org/10.1016/S0048-7333(99)00041-4

Madhavan, R., & Prescott, J. (2017). *Chapter 20: The network perspective of alliances: taking stock and looking ahead*. Edward Elgar Publishing. https://doi.org/10.4337/9781783479580.00033

Martin, B. R. (2016). R&D policy instruments – a critical review of what we do and don't know. *Industry and Innovation*, *23*(2), 157–176. https://doi.org/10.1080/13662716.2016.1146125

Martínez-Noya, A., & Narula, R. (2018). What more can we learn from R&D alliances? A review and research agenda. *BRQ Business Research Quarterly*, *21*, 195–212. http://creativecommons.org/licenses/by-nc-nd/4.0

Masciarelli, F., Laursen, K., & Prencipe, A. (2010). Trapped by over-embeddedness: the effects of regional social capital on internationalization. *DRUID Working Paper Series*.

McCarthy, S. (2017). Success Rates in Horizon 2020. *Journal of Innovation Management*, *5*(3), 18–22. https://doi.org/https://doi.org/10.24840/2183-0606_005.004_0003

Morgan, K. (1997). The Learning Region: Institutions, Innovation and Regional Renewal. *Regional Studies*, *31*(5), 491–503. https://doi.org/10.1080/00343409750132289

Morrison, A. (2008). Gatekeepers of knowledge within industrial districts: Who they are, how they interact. *Regional Studies*, *42*(6), 817–835. https://doi.org/10.1080/00343400701654178

Muller, E., & Zenker, A. (2001). Business services as actors of knowledge transformation: the role of KIBS in regional and national innovation systems. *Research Policy*, *30*(9), 1501–1516. https://doi.org/https://doi.org/10.1016/S0048-7333(01)00164-0

Murphy, K. R., & Aguinis, H. (2022). Reporting Interaction Effects: Visualization, Effect Size, and Interpretation. *Journal of Management*, *48*(8), 2159–2166. https://doi.org/10.1177/01492063221088516

Muscio, A., & Vallanti, G. (2014). Perceived Obstacles to University–Industry Collaboration: Results from a Qualitative Survey of Italian Academic Departments. *Industry and Innovation*, *21*(5), 410–429. https://doi.org/10.1080/13662716.2014.969935

Nelson, R. R. (2004). The market economy, and the scientific commons. *Research Policy*, *33*(3), 455–471. https://doi.org/https://doi.org/10.1016/j.respol.2003.09.008

Nooteboom, B. (2000). Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. *Journal of Management and Governance*, *4*(1), 69–92. https://doi.org/10.1023/A:1009941416749

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

OECD. (1997). *National Innovation Systems*. https://doi.org/10.1007/978-3-319-15347-6_458

Oughton, C., Landabaso, M., & Morgan, K. (2002). The regional innovation paradox: Innovation policy and industrial policy. *Journal of Technology Transfer*, *27*(1), 97–110. https://doi.org/10.1023/A:1013104805703

Owen-Smith, J., & Powell, W. W. (2004). Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community. *Organization Science*, *15*(1), 5–21. https://doi.org/10.1287/orsc.1030.0054

Paier, M., & Scherngell, T. (2011). Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model. *Industry and Innovation*, *18*(1), 89–104. https://doi.org/10.1080/13662716.2010.528935

Phelps, C., Heidl, R., & Wadhwa, A. (2012). Knowledge, Networks, and Knowledge Networks: A Review and Research Agenda. *Journal of Management*, *38*(4), 1115–1166. https://doi.org/10.1177/0149206311432640

Podolny, J. M. (2001). Networks as the Pipes and Prisms of the Market. *American Journal of Sociology*, *107*(1), 33–60. https://doi.org/10.1086/323038

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

Pontikakis, D., González Vázquez, I., Bianchi, G., Ranga, M., Marques Santos, A., Reimeris, R., Mifsud, S., Morgan, K., Madrid, C., & Stierna, J. (2022). *Partnerships for Regional Innovation – Playbook*. Publications Office of the European Union. https://doi.org/10.2760/775610

Porter, M. E. (2000). Location, Competition, and Economic Development: Local Clusters in a Global Economy. *Economic Development Quarterly*, *14*(1), 15–34. https://doi.org/10.1177/089124240001400105

Quintane, E., & Carnabuci, G. (2016). How do brokers broker? Tertius gaudens, tertius iungens, and the temporality of structural holes. *Organization Science*, *27*(6), 1343–1360. https://doi.org/10.1287/orsc.2016.1091

Reagans, R., & McEvily, B. (2003). Network Structure and Knowledge Transfer: The Effects of Cohesion and Range. *Administrative Science Quarterly*, *48*(2). https://doi.org/10.2307/3556658

Reuer, J. J., & Devarakonda, R. (2017). Partner Selection in R&D Collaborations: Effects of Affiliations with Venture Capitalists. *Organization Science*, *28*(3), 574–595. https://doi.org/10.1287/orsc.2017.1124

Revet, K. (2022). *Social networks and knowledge creation: A relational approach to resources flowing in scientific collaboration networks*. Grenoble Ecole de Management.

Roediger-Schluga, T., & Barber, M. J. (2006). *The structure of R&D collaboration networks in the European Framework Programmes* (MERIT Working Papers, Issues 2006–036). United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT). https://econpapers.repec.org/RePEc:unm:unumer:2006036

Rossi, F., Caloffi, A., & Russo, M. (2016). Networked by design: Can policy requirements influence organisations’ networking behaviour? *Technological Forecasting and Social Change*, *105*, 203–214. https://doi.org/10.1016/j.techfore.2016.01.004

Rothaermel, F. T., & Boeker, W. (2008). Old technology meets new technology: complementarities, similarities, and alliance formation. *Strategic Management Journal*, *29*(1), 47–77. https://doi.org/https://doi.org/10.1002/smj.634

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

Sammut, C., & Webb, G. I. (Eds.). (2017). *TF–IDF BT - Encyclopedia of Machine Learning and Data Mining* (p. 1274). Springer US. https://doi.org/10.1007/978-1-4899-7687-1_832

Santoro, M. D., & Saparito, P. A. (2003). The firm's trust in its university partner as a key mediator in advancing knowledge and new technologies. *IEEE Transactions on Engineering Management*, *50*(3), 362–373. https://doi.org/10.1109/TEM.2003.817287

Scherngell, Thomas, & Barber, M. J. (2011). Distinct spatial characteristics of industrial and public research collaborations: Evidence from the fifth EU Framework Programme. *Annals of Regional Science*, *46*(2), 247–266. https://doi.org/10.1007/s00168-009-0334-3

Schilling, M. A., & Phelps, C. C. (2007). Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation. *Management Science*, *53*(7), 1113–1126. https://doi.org/10.1287/mnsc.1060.0624

Shibayama, S., Walsh, J. P., & Baba, Y. (2012). Academic Entrepreneurship and Exchange of Scientific Resources: Material Transfer in Life and Materials Sciences in Japanese Universities. *American Sociological Review*, *77*(5), 804–830. https://doi.org/10.1177/0003122412452874

Simensen, E. O., & Abbasiharofteh, M. (2022). Sectoral patterns of collaborative tie formation: investigating geographic, cognitive, and technological dimensions. *Industrial and Corporate Change*, *31*(5), 1223–1258. https://doi.org/10.1093/icc/dtac021

Simmie, J. (2011). 40. Learning regions. In P. Cooke, B. Asheim, R. Boschma, R. Martin, D. Schwartz, & F. Tödtling (Eds.), *Handbook of Regional Innovation and Growth* (pp. 547–556). Edward Elgar Publishing Limited.

Singh, J. (2005). Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science*, *51*(5), 756–770. http://www.jstor.org/stable/20110371

Smiljic, S. (2020). Beyond the dyad: Role of non-competitive partners in coopetitive R&D projects. *International Journal of Innovation Management*, *24*(8), 1–25. https://doi.org/10.1142/S136391962040006X

Snijders, T. A. B. (1999). Prologue to the Measurement of Social Capital. *The Tocqueville Review*, *20*(1), 27–44. https://doi.org/10.3138/ttr.20.1.27

Sorenson, O., Rivkin, J. W., & Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy*, *35*(7), 994–1017. https://doi.org/10.1016/j.respol.2006.05.002

Stuart, T. E. (2000). Interorganizational alliances and the performance of firms: a study of growth and innovation rates in a high-technology industry. *Strategic Management Journal*, *21*(8), 791–811. https://doi.org/https://doi.org/10.1002/1097-0266(200008)21:8<791::AID-SMJ121>3.0.CO;2-K

Takalo, T., Tanayama, T., & Toivanen, O. (2013). Market failures and the additionality effects of public support to private R&D: Theory and empirical implications. *International Journal of Industrial Organization*, *31*(5), 634–642. https://doi.org/https://doi.org/10.1016/j.ijindorg.2013.02.002

Ter Wal, A. L. J. (2014). The dynamics of the inventor network in german biotechnology: Geographic proximity versus triadic closure. *Journal of Economic Geography*, *14*(3), 589–620. https://doi.org/10.1093/jeg/lbs063

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

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

Tóth, G., Juhász, S., Elekes, Z., & Lengyel, B. (2021). Repeated collaboration of inventors across European regions. *European Planning Studies*, *29*(12), 2252–2272. https://doi.org/10.1080/09654313.2021.1914555

Tsouri, M. (2019). Knowledge transfer in time of crisis: evidence from the Trentino region. *Industry and Innovation*, *26*(7), 820–842. https://doi.org/10.1080/13662716.2018.1551124

Tsouri, M. (2022). Knowledge networks and strong tie creation: the role of relative network position. *Journal of Geographical Systems*, *24*(1), 95–114. https://doi.org/10.1007/s10109-021-00351-9

Tsouri, M., & Pegoretti, G. (2020). Structure and resilience of local knowledge networks: the case of the ICT network in Trentino. *Industry and Innovation*, *00*(00), 1–20. https://doi.org/10.1080/13662716.2020.1775070

Usai, S., Marrocu, E., & Paci, R. (2017). Networks, Proximities, and Interfirm Knowledge Exchanges. *International Regional Science Review*, *40*(4), 377–404. https://doi.org/10.1177/0160017615576079

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, *61*(4), 674–698. https://doi.org/10.4324/9780429494338

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, *111*(2), 447–504. https://doi.org/10.1086/432782

Vicente, J., Balland, P. A., & Brossard, O. (2011). Getting into networks and clusters: Evidence from the Midi-Pyrenean global navigation satellite systems (GNSS) collaboration network. *Regional Studies*, *45*(8), 1059–1078. https://doi.org/10.1080/00343401003713340

Walker, G., Kogut, B., & Shan, W. (1997). Social Capital, Structural Holes and the Formation of an Industry Network. *Organization Science*, *8*(2), 109–125. https://doi.org/10.1287/orsc.8.2.109

Wanzenböck, I., Lata, R., & Ince, D. (2020). Proposal success in Horizon 2020: A study of the influence of consortium characteristics. *Quantitative Science Studies*, *1*(3), 1136–1158. https://doi.org/10.1162/qss_a_00067

Werker, C., Korzinov, V., & Cunningham, S. (2019). Formation and output of collaborations: the role of proximity in German nanotechnology. *Journal of Evolutionary Economics*, *29*(2), 697–719. https://doi.org/10.1007/s00191-019-00605-2

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science (New York, N.Y.)*, *316*(5827), 1036–1039. https://doi.org/10.1126/science.1136099

Wuyts, S., Colombo, M. G., Dutta, S., & Nooteboom, B. (2005). Empirical tests of optimal cognitive distance. *Journal of Economic Behavior and Organization*, *58*(2), 277–302. https://doi.org/10.1016/j.jebo.2004.03.019

Xavier Molina-Morales, F., Belso-Martínez, J. A., Más-Verdú, F., & Martínez-Cháfer, L. (2015). Formation and dissolution of inter-firm linkages in lengthy and stable networks in clusters. *Journal of Business Research*, *68*(7), 1557–1562. https://doi.org/https://doi.org/10.1016/j.jbusres.2015.01.051

Zaheer, A., McEvily, B., & Perrone, V. (1998). Does Trust Matter? Exploring the Effects of Interorganizational and Interpersonal Trust on Performance. *Organization Science*, *9*(2), 141–159. https://doi.org/10.1287/orsc.9.2.141

Appendices

Appendix 2.A. A full list of NACE Rev.2 division categories of economic activity, represented in our dataset, alongside the number of firms which belong to each category in our sample.

NACE div.	Description of the two-digit NACE Rev.2 division	Num of firms
71	Architectural and engineering activities; technical testing and analysis	52
62	Computer programming, consultancy and related activities	49
22	Manufacture of rubber and plastic products	36
13	Manufacture of textiles	32
46	Wholesale trade, except of motor vehicles and motorcycles	31
72	Scientific research and development	31
20	Manufacture of chemicals and chemical products	28
28	Manufacture of machinery and equipment n.e.c	20
25	Manufacture of fabricated metal products, except machinery equipment	16
23	Manufacture of other non-metallic mineral products	14
86	Human health activities	11
36	Water collection, treatment and supply	10
74	Other professional, scientific and technical activities	10
10	Manufacture of food products	9
42	Civil engineering	9
43	Specialized construction activities	9
52	Warehousing and support activities for transportation	9
26	Manufacture of computer, electronic and optical products	7
27	Manufacture of electrical equipment	7
29	Manufacture of motor vehicles, trailers and semi-trailers	6
38	Waste collection, treatment and disposal activities; materials recovery	6
49	Land transport and transport via pipelines	6
82	Office administrative, office support and other business support activities	6
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations	5
32	Other manufacturing	5
61	Telecommunications	5
70	Activities of head offices; management consultancy activities	5
81	Services to buildings and landscape activities	5
01	Crop and animal production, hunting and related service activities	4
47	Retail trade, except of motor vehicles and motorcycles	4
15	Manufacture of leather and related products	3
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	3
17	Manufacture of paper and paper products	3
30	Manufacture of other transport equipment	3
31	Manufacture of furniture	3
35	Electricity, gas, steam and air conditioning supply	3
63	Information service activities	3

73	Advertising and market research	3
80	Security and investigation activities	3
11	Manufacture of beverages	2
24	Manufacture of basic metals	2
33	Repair and installation of machinery and equipment	2
41	Construction of buildings	2
68	Real estate activities	2
87	Residential care activities	2
08	Other mining and quarrying	1
14	Manufacture of wearing apparel	1
37	Sewerage	1
45	Wholesale and retail trade and repair of motor vehicles and motorcycles	1
50	Water transport	1
55	Accommodation	1
64	Financial service activities, except insurance and pension funding	1
85	Education	1
88	Social work activities without accommodation	1
91	Libraries, archives, museums and other cultural activities	1
95	Repair of computers and personal and household goods	1
96	Other personal service activities	1
	Total number of firms in the sample:	498

Appendix 3.A. Results of Quadratic Assignment Procedure regression (replications: 100).

	Dependent variable: Partner t+1							
	(Model 1.0)		(Model 1.1)		(Model 1.2)		(Model 1.3)	
	Estimate	Exp(b)	Estimate	Exp(b)	Estimate	Exp(b)	Estimate	Exp(b)
CollabExp			**1.888**	6.603				
FailedExp					**3.878**	48.309	**3.915**	50.127
SuccessExp							**1.810**	6.109
MixedExp							**3.341**	28.252
GeoDist	**-0.002**	0.998	**-0.001**	0.999	**-0.001**	0.999	**-0.001**	0.999
ExpDiff	**0.140**	1.151	**0.138**	1.148	**0.137**	1.147	**0.134**	1.144
TC	**0.415**	1.516	**0.463**	1.589	**0.618**	1.855	**0.677**	1.968
PRO	**2.016**	7.510	**2.007**	7.440	**1.877**	6.532	**1.854**	6.391
LargeFirm	**0.897**	2.452	**0.880**	2.411	**0.961**	2.615	**0.943**	2.567
Constant	**-6.605**	0.001	**-6.613**	0.001	**-6.686**	0.001	**-6.696**	0.001
Pseudo R2 measures:								
(Dn-Dr)/ (Dn-Dr+dfn)	0.575		0.575		0.575		0.575	
(Dn-Dr)/Dn	0.976		0.977		0.977		0.977	
Fraction predicted 1s correct	0.375		0.453		0.456		0.454	
Fraction predicted 0s correct	0.997		0.997		0.997		0.997	

Note: Values in bold are statistically significant at the 1%.

Appendix 3.B. Variables correlation matrix (firm-firm subsample).

Parameter	SuccessExp	FailedExp	MixedExp	CogDist	GeoDist	ExpDiff	TC	LargeFirm	AgeDiff
Partner	0.019	0.093	0.000	-0.027	**-0.023**	0.003	-0.001	**0.016**	0.006
SuccessExp		-0.001	0.000	-0.015	**-0.016**	**0.013**	-0.004	**0.012**	0.004
FailedExp			0.000	-0.021	**-0.018**	-0.001	-0.003	-0.008	-0.005
MixedExp				0.002	-0.003	0.007	-0.001	0.005	0.000
CogDist					**0.046**	0.002	**-0.016**	**0.040**	**0.056**
GeoDist						**-0.059**	**-0.050**	**-0.017**	**0.026**
ExpDiff							**0.033**	**0.103**	**0.051**
TC								**0.015**	**0.014**
LargeFirm									**0.204**

Note: Values in bold are statistically significant ($p < 0.01$).