Sapienza Università di Roma

Sapienza University of Rome

Department of Social Sciences and Economics
PhD in Markets and Institutions for Development and Innovation

Thesis For The Degree Of Doctor Of Philosophy

A network-based approach to the analysis of entrepreneurial ecosystems and collaborative R&D

Supervisors
Prof. Roy Cerqueti
Prof. Gianluca Vagnani

Candidate
Andrea Ancona

Academic year 2023/2024 - XXXVI cycle

Preface

In July 2020, I graduated in Management Engineering. Until that moment, I had not seriously considered the possibility of pursuing an academic career. My background predominantly focused on hard disciplines, and I envisioned my future immersed in "hands-on" activities related to the management of a company's supply chain, production, and operations. However, as it happened for many others in that period, a breakthrough event shocked our world, serving as a significant turning point in our lives. I should have started an internship in March 2020, but the outbreak of the pandemic halted most in-person activities. It was at that stage that I decided to pivot towards a research-driven thesis for my master's degree, focusing on the application of Social Network Analysis and complexity theory to an entrepreneurial case study. Day by day, my interest in doing research burgeoned, fueled by unexpected curiosity, passion, and enthusiasm for a path that was so far from my initial plans.

More than three years later, I am concluding my PhD, and I could not have been more satisfied with the choice I made. Nevertheless, this path was far from easy. This thesis encompasses years of hard work, frustration, anxiety, and rejections. Doing a PhD, especially in Italy, is a kind of suspended period in your life. Some perceive you as a student; a few as a worker; the majority fail to comprehend your endeavors. It is a roller-coaster, where moments of doubt and uncertainty abound. You can feel isolated, confined within a closed community, with limited possibility of reaching out to the rest of the world. Yet, how much satisfaction derives from seeing your name associated with an academic article, a scientific presentation, or an event organization? How rewarding is it to present your own work in front of your community, unveiling parts of your vision and identity each time you have the chance to do it?

From the outside, a PhD journey is often perceived as an extremely narrow experience, providing expertise on specific topics. However, what we actually learn extends far beyond that. Writing and communications skills, critical thinking, time and project management, resilience, and adaptability to pressing circumstances - everything driven by curiosity and willingness to learn - are just a few of the aspects you have the chance to develop along this path.

Undoubtedly, as PhD candidates, we often feel lost about what we are doing. We (reasonably) want our research to be more impactful and concrete. But, in the end, is there anything more real and concrete than producing, exchanging, and disseminating knowledge?

Contents

Introduction

The emerging paradigms of open innovation (Chesbrough, 2003) have increasingly determined the need to collaborate for generating and disseminating knowledge, as well as developing technological innovation. Nowadays, innovation rarely occurs in isolation. Research and Development (R&D) investments frequently include the presence of partners, producing a growing interest in collaborative R&D. Several researchers have delved into the concept of collaborative R&D, investigating the determinants behind the choice of partners (Diestre and Rajagopalan, 2012; Reuer and Devarakonda, 2017), the effects of collaboration on innovation performance (Hoang and Rothaermel, 2005, 2010), the dynamics of collaborative relationships (Tatarynowicz et al., 2016; Jakobsen et al., 2019), and many other facets of this phenomenon. At the same time, the increasing interconnectedness among firms, governments, entrepreneurs, research institutes, and universities has prompted scholars to adopt an ecosystem approach in analyzing collaborative and innovation processes. This approach indeed, enables the simultaneous investigation of social, cultural, economic, and technological factors affecting the entrepreneurial process (Spigel, 2017; Stam and Van de Ven, 2021).

In this context, the notion of economic complexity has gained relevance in the literature. The adoption of this perspective allows for the reduction of the computational dimension of socio-economic phenomena, while preserving more information than traditional approaches. In particular, methods from spectral analysis and network theory effectively capture the complex nature of socio-economic processes, offering tools for their systemic analysis. Recently, various institutions, such as the World Bank, the European Commission, the World Economic Forum, and the OECD, have embraced the paradigm of economic complexity to develop more tailored and data-driven policies (Balland et al., 2022). Also in the entrepreneurship literature, the complexity lens has started to be adopted for providing a detailed picture of the entrepreneurial process, supporting and strengthening the implementation of an ecosystem approach (Mazzoni et al., 2021).

Within the realm of economic complexity, economic networks play a significant role. The origins of research on economic networks can be traced back to two distinct disciplines: sociology and physics. The use of networks in the economic field is crucial to analyze the dynamic interactions among groups of heterogeneous actors, providing insights into their systemic behavior, and the link with the underlying structural network properties. More specifically, studies on social, technological, and financial networks have proliferated, with a primary focus on exploring the interactions between different entities in the form of knowledge and technological transfers, R&D alliances, credit exchange, trades, and many other types of relationships (Schweitzer et al., 2009).

This thesis aims to contribute to this strand of research by employing network theory with a precise focus on collaborative environments and innovation systems. Collaborative networks are indeed a popular concept in innovation and R&D literature, and they have been studied for various purposes, including their role in achieving a higher degree of novelty in product innovation (Nieto and Santamaría, 2007), the impact of absorptive capacity among different types of partners on product innovation performance (Tsai, 2009), the relevance of different collaboration strategies to facilitate knowledge diffusion (Campos et al., 2013), and the effect of R&D subsidies on the total welfare of alliances (König et al., 2019). According to this representation, distinct organizations correspond to the network nodes, while edges stand for collaborative relationships between them.

Thus, the idea behind this research is that a network science approach can unveil behavioral patterns within innovation systems, and contribute to shedding light on the interaction dynamics characterizing inter-organizational contexts (Najafian and Colabi, 2014; Ferraro and Iovanella, 2017). Network science involves the study of the theoretical foundations of network structure and dynamic behaviors, and it finds application in various subfields where there is an interest in exploring complex properties such as structure, dynamism, bottom-up evolution, autonomy, stability, and emergence (Lewis, 2009). Many real systems exhibit a complex nature in which the interactions among constituent elements determine the collective behavior of the system as a whole (Funtowicz and Ravetz, 1994; Thurner et al., 2018). In more detail, complex systems are composed of networks of interacting entities possessing specific structural features (Barabási, 2016; Newman, 2018; Russell and Smorodinskaya, 2018), along with non-linear and heterogeneous interactions (Thurner et al., 2018). These characteristics make networks the most suitable representation for analyzing their structure, roles, and dynamics. In particular, Social Network Analysis (SNA) is one of the most popular methods to examine relationships among social groups and entities through the analysis of their connection structure (Hu et al., 2015) and the computation of centrality measures (Scott and Carrington, 2011).

SNA has its foundations in graph theory, a discipline whose birth is attributed to the resolution of the famous "Könisberg Bridge Problem" (West et al., 2001). In general terms, a graph G is described by a couple (V, E), where V is the set of vertices (or nodes), and E is the set of edges (or links). Relationships between any pairs of nodes are described mathematically by a $n-$square matrix A, called adjacency matrix. An element a_{ij} of A is equal to 1 if node i is adjacent (i.e., linked) to node j, and 0 otherwise. The degree of node i is equal to the number of its adjacent nodes, and it is denoted as k_i. Alternatively, let us assume that G is a weighted graph, i.e., its edges are weighted based on factors such as the frequency of occurrence of the associated relations. In this case, G is described by a weighted adjacency matrix W, whose elements are equal to the weight w_{ij} if $(i, j) \in E$, and 0 otherwise. Thus, a network is ultimately represented by a graph, in which nodes correspond to actors and edges stand for specific directed or undirected relationships between them. By representing social interactions through the use of networks, we are able to evaluate structural network properties quantitatively. To this aim, a series of network metrics and approaches will be introduced and widely described within the different chapters of this thesis.

The thesis is organized as follows. The first chapter investigates the concept of entrepreneurial ecosystem through the lens of network theory. In particular, it contributes to the existing body of literature by providing insights into the structural features of such systems (Wurth et al., 2022). The acknowledged complexity of entrepreneurial ecosystems indeed, generates a growing demand for transdisciplinary works that introduce methods and techniques from other research fields, such as simulation and statistical approaches, agent-based modeling, and SNA (Carayannis et al., 2022). The main outcome of the first chapter is the formulation of seven network-based principles associating specific network metrics with distinct structural features of entrepreneurial ecosystems. These metrics are aimed to support the measurement of an entrepreneurial ecosystem and the design of policy interventions in case of unmet properties. The proposed methodology is then applied to an original network of European start-ups on Twitter. This case study represents a further contribution to the field by presenting a novel way to conceptualize entrepreneurial ecosystems, considering online interactions among actors. The results suggest a partial ecosystem-like nature of the analyzed network, providing evidence about possible policy recommendations.

The second chapter is devoted to the study of different partner selection mechanisms in collaborative research networks. In particular, it contributes to the collaborative R&D literature by exploring the effect of collaborating with new over existing partners on the amount of funding received from the European Commission within the eighth EU Framework Programme (FP), Horizon 2020 (H2020). Specifically, this chapter employs an innovative technique in SNA, known as the "dual-projection" approach (Everett and Borgatti, 2013). This method is specifically designed for analyzing two-mode social network data, i.e., data characterizing bipartite networks. Collaborative research network data belongs to this category, being constituted by two distinct sets of nodes (i.e., organizations and projects), with connections occurring only between nodes belonging to different sets, representing the participation of organizations in projects. The results from the econometric analysis show that partnering with new organizations increases the probability of obtaining more funds when entering a new project compared to collaborating with already existing partners. Additionally, projects coordinated by private or public organizations appear more likely to secure higher funding compared to research centers and higher education institutions. However, the relevance of partners' connections decreases as we move away from the focal organization. These findings provide valuable insights into the effectiveness of the European Research and Technological Development Policy.

Finally, the third chapter focuses on the analysis of collaboration patterns and participation dynamics in projects funded by the first eight EU FPs, from FP1 to H2020. It combines elements from SNA and statistics through the lens of economic complexity by computing participants' centrality and assessing the stochasticity of the process, further estimating the probability of transitioning between classes of centrality across consecutive FPs. In doing so, this chapter contributes to the literature on the dynamics of collaborative research projects, particularly EU-funded ones, by shedding light on the Markovian nature of the collaboration dynamics over an extended period. Most of the existing works indeed, employ Stochastic Actor-Oriented Models (SAOMs) to explore the mechanisms driving network evolution (Giuliani, 2013; Cao et al., 2017) since they are efficient in processing longitudinal network data (Broekel et al., 2014). Nevertheless, SAOMs assume that network structures evolve as a Markov chain outcome (Snijders, 2017), whereas in many real networks, evolution often displays a non-Markovian behavior (Williams et al., 2022). The results reveal a quasi-Markovianity of participation dynamics, opening up opportunities for accurate forecasting procedures to estimate the leaders of future FPs. "Preferential attachment" mechanisms emerge, thus confirming the relevance of participating in EU-funded projects to strengthen organizations' popularity. On the other hand, this finding highlights the pressing need to address "oligopolistic" behaviors linked to European funds that hinder the full realization of the European Research Area (ERA). The crucial role of European policies is emphasized by the estimated transition probabilities, which are influenced by breakthrough events in the EU research framework like the Treaty of Maastricht and the promotion of the ERA. However, sustained efforts are necessary to ensure a certain degree of openness and the "democratization" of European research funds.

It is important to specify that all the analyses throughout the entire thesis have been conducted in R (R Core Team, 2014).

Chapter 1

Network-based principles of entrepreneurial ecosystems: a case study of a start-up network

Abstract

Entrepreneurial ecosystems are wealthy environments in which entrepreneurs, firms, and governments can operate frictionless, contributing to innovation and economic growth. The investigation of the structure of such systems is an open issue. We provide insights on this aspect through the formulation of seven network-based principles associating specific network metrics to distinct structural features of entrepreneurial ecosystems. In this way, we aim to support the measurement of the structural characteristics of an entrepreneurial ecosystem and the design of policy interventions in case of unmet properties. The proposed methodology is applied to an original network built on the relationships occurring on Twitter among 612 noteworthy start-ups from seven different European countries. This is a novel way to conceptualize entrepreneurial ecosystems considering online interactions. Thus, this work represents a first attempt to analyze the structure of entrepreneurial ecosystems considering their network architecture to guide policy-making decisions. Our results suggest a partial ecosystem-like nature of the analyzed network, providing evidence about possible policy recommendations.

Keywords: Entrepreneurial ecosystem, Network-based approach, Network-based principles, Complexity theory, Start-ups, Social media.

1.1 Introduction

The concept of ecosystem is an integral part of the entrepreneurial world. It was borrowed from the field of ecology during the late 80's (Van de Ven, 1986), where ecosystems are characterized by strongly interconnected groups of organisms that interact with each other and with the physical environment where they live, in order to maintain the dynamic equilibrium of the system (Jackson, 2011). The ecosystem metaphor marked a substantial shift in entrepreneurship studies, away from an individualistic perspective toward a deeper attention to collectivity and community, thus including social, cultural and economic forces in the entrepreneurship process (Stam and Van de Ven, 2021). However, despite many years passed since its first introduction, the concept of entrepreneurial ecosystem gained momentum only recently (Stam, 2015), especially in the fields of management and innovation, as the growing number of papers published after 2010 proves (Tsujimoto et al., 2018). One of the possible reasons behind such a growing popularity may be that the ecosystem metaphor is one of the most apt for representing the many facets of active entrepreneurial environments.

The identification and the measurement of entrepreneurial ecosystems is crucial in the innovation field since actors, structures, and relations that characterize ecosystems lead to digital innovation and vice versa (Hinings et al., 2018; Granstrand and Holgersson, 2020). Nevertheless, the investigation of the structure of entrepreneurial ecosystems is an open issue in the current literature, as highlighted by Wurth et al. (2022) in their proposal for an "Entrepreneurial Ecosystem Research Program". Our paper aims to shed light on this aspect, considering the architecture of a system a key factor for the diffusion of innovation as well as a key determinant for its global performance (Ferraro and Iovanella, 2016; Muller and Peres, 2019).

In more detail, we propose to contribute to the literature by adopting an innovative approach to the measurement of entrepreneurial ecosystems with respect to other methods, such as the European Index of Digital Entrepreneurship Systems (EIDES) family of indicators (Autio et al., 2020). According to Roundy et al. (2018) and Leendertse et al. (2021) indeed, entrepreneurial ecosystems can be analyzed using complexity theory to understand better their internal dynamics and the interactions among their members. In this way, we aim at answering two different research questions. Firstly, "what are the elements highlighted in the literature that can be identified as structural features of an entrepreneurial ecosystem?". Secondly, looking at these elements from a complex systems perspective, "is it possible to quantify such elements by means of network metrics?". Indeed, the most common way to represent complex systems is through the use of networks (Thurner et al., 2018). Moreover, according to Seppä and Tanev (2011), the presence of a network structure is a necessary condition for the existence of an ecosystem. Then, networks constitute valid, apt and available models to represent entrepreneurial ecosystems.

We address the aforementioned research questions by formulating seven principles, each of them referring to a structural feature of entrepreneurial ecosystems that can be expressed by means of a consolidated metric from network science theory. Therefore, the first contribution of this study is the proposal of seven network-based metrics that aim to be a support to the measurement of the structural characteristics of an entrepreneurial ecosystem. In doing so, we propose to provide a set of indicators that may help several actors - such as practitioners and policy makers - monitor the structure of an entrepreneurial ecosystem and design policy interventions aimed at fulfilling eventual missing properties.

Secondly, another key contribution of this paper is represented by measuring entrepreneurial ecosystems via social media data. Considering online connections to map entrepreneurial ecosystems is innovative to this field. According to the literature, entrepreneurial ecosystems are locally defined, and their actors interact mainly through social, financial, technological, and economic flows (Spigel, 2017; Stam and Van de Ven, 2021). Under this perspective, proximity emerges as a critical factor and enabler of entrepreneurial ecosystems (Isenberg, 2011). In this vein, analyzing ecosystems through online relationships contributes to extending the concept of proximity beyond regional and national borders, allowing the conceptualization of entrepreneurial ecosystems at the continental level. In fact, we report a case study of a Twitter network constituted by some among the most popular and promising start-ups belonging to seven different countries: Austria, France, Germany, Italy, Spain, Switzerland, and the United Kingdom. Their names are collected from accredited sources which provide every year a list of the most noteworthy start-ups in the international context.

Being start-ups the embodiment of entrepreneurship and innovation (Feldman, 2001; Isenberg, 2016), a start-up network seems a well-suited candidate for testing the characteristics of an entrepreneurial ecosystem. In fact, according to the widely used definition of Blank (2010), "a start-up is an organization formed to search for a repeatable and scalable business model". Furthermore, the definition of a start-up ecosystem as "a society of founders with ideas and skills, young companies at early stages with talent, incubators with mentors and capital, early adopters, and media" (Aleisa, 2013) is definitely included in the wider concept of entrepreneurial ecosystem. Then, for all the actors of a start-up network it is important to understand how the structure of this system affects their own activities (Spender et al., 2017).

The network that we take into account is built on an original dataset obtained by mining relationships among 612 start-ups on Twitter, a social networking platform frequently used for testing professional links across start-ups (Gloor et al., 2013; Perotti and Yu, 2015; Yue et al., 2019; Tumasjan et al., 2021). In the choice of Twitter as a proxy to build the network, we also rely on the study of Tajudeen et al. (2018), which sustain that well performing innovative organizations show a high social media usage, and on the paper of Olanrewaju et al. (2020), that provide a comprehensive review on the importance of social media for entrepreneurship.

The paper is organized as follows. An overview of the state of the art on entrepreneurial ecosystems is reported in Section 1.2. Section 1.3 describes the approach of this study, which has its foundations in Social Network Analysis and complexity theory, before formulating the network-based principles characterizing the structure of entrepreneurial ecosystems. Section 1.4 introduces the case study, including the processes of data collection and network construction. The results obtained by the measurement of the network-based principles are reported in Section 1.5. Finally, Section 1.6 points out the main theoretical and practical contributions of this study, as well as the acknowledged possible limitations and further developments.

1.2 The concept of entrepreneurial ecosystem

In the last decade, the number of publications on entrepreneurial ecosystems has dramatically increased (Cao and Shi, 2021), and, more generally, the adoption of the ecosystem concept in entrepreneurship and innovation studies has become increasingly popular (Tsujimoto et al., 2018). The origins of such a phenomenon in the business and management fields can be traced back to the end of the 1980s and the early 1990s, thanks to the pioneering works of Van de Ven (1986) and Moore (1993). Stemming from these papers, several concepts emerged, including business, innovation, digital, knowledge, and entrepreneurial ecosystems. Despite sharing the common definition of ecosystems, these constructs differentiate from each other for various reasons. In particular, one of the main differences between entrepreneurial ecosystems and business or innovation ecosystems is that while the performance of the latter depend on the strategies of a leading organization (or the orchestrator), the emergence and health of entrepreneurial ecosystems are strictly related to regional development, thus exhibiting a strong connection with economic growth (Hakala et al., 2020).

The entrepreneurial ecosystem approach finds its roots in the literature on regional development and strategic management. Indeed, there is a notable interest in exploring the socioeconomic characteristics and performance of innovation systems, as in the regional development literature. At the same time, studies on entrepreneurial ecosystems frequently adopt a value creation perspective, typical of the strategic management literature (Acs et al., 2017). Nevertheless, entrepreneurial ecosystems distinguish themselves from traditional clusters in their emphasis on leveraging digital affordances, evolving around entrepreneurial opportunity discovery and pursuit, prioritizing business model innovation, fostering voluntary horizontal knowledge spillovers, and identifying entrepreneurial opportunities externally (Autio et al., 2018). Another significant distinction lies in the boundaries of clusters and entrepreneurial ecosystems. While the former are bounded by a specific industry, the latter develop at a higher level, encompassing multiple industries, as well as entrepreneurial agents and institutions (Auerswald and Dani, 2017). Moreover, the role of governments and the ease to access regional resources contribute to the differentiation of entrepreneurial ecosystems from traditional clusters and regional innovation systems (Spigel and Harrison, 2018).

Even though authors do not converge toward a specific definition of entrepreneurial ecosystems, the vast majority of them agree in characterizing entrepreneurial ecosystems as multifaceted structures (Van de Ven, 1993; Spigel, 2017), with outputs generated from the combination of financial, knowledge, institutional, and social capital (Nicotra et al., 2018). Such complexity is inherent in the concept of ecosystem, which, in the case of entrepreneurial ecosystems, indicates a variety of components closely related to the economic, technological, and societal context (Audretsch et al., 2019). In particular, entrepreneurial ecosystems are characterized by networks of knowledge, labor, and social capital (Malecki, 2018; Greve and Salaff, 2003), whose innovation outcomes are influenced by various factors, including industrial and technological elements, organizational aspects, institutional and policy frameworks, as well as social, temporal, and spatial characteristics (Autio et al., 2014).

Emphasis is given to the interactions among the members of the ecosystem, that should result into high network density, many connecting events, collaboration between large companies and local start-ups, together with an easy access to all kinds of relevant resources, and an enabling role of government in the background (Feld, 2012). Nevertheless, without some degrees of cohesion and shared values among the agents of an entrepreneurial ecosystem, they will operate autonomously, without commonality in their activities, and differently with respect to a system of interrelated actors (Roundy, 2017). Additionally, an entrepreneurial-market logic is essential, encapsulating a related set of goals and behaviors focused on innovation, the creation of new markets, new business models and technologies, and tolerance for uncertainty and failure (Cunningham et al., 2002). Spigel (2022) integrates previous contributions, finding that entrepreneurial ecosystems are nested rather than cohesive.

One of the main aspects that emerges in many works is the fundamental role that governments, and more specifically the policies they adopt, have for the correct establishment of an entrepreneurial ecosystem. Governments' policies shape the institutional environment in which entrepreneurial decisions are made, by leveraging the presence of local research centers, increasing the availability of venture capital, encouraging a culture of risk taking, and creating strong local informational and business development networks (Feldman, 2001; Minniti, 2008). However, policy initiatives to promote the growth of innovative companies depend on the specific stage of firms development (Audretsch et al., 2020).

Several authors describe entrepreneurial ecosystems as complex systems, thus borrowing concepts from complexity theory. One of the most comprehensive contributions in this regard comes from Roundy et al. (2018), who conceptualize entrepreneurial ecosystems as complex adaptive systems by providing the following definition: "An entrepreneurial ecosystem is a self-organized, adaptive, and geographically bounded community of complex agents operating at multiple, aggregated levels, whose non-linear interactions result in the patterns of activities through which new ventures form and dissolve over time." The adoption of the ecosystem metaphor indeed, implicitly assumes the existence of a self-organizing, self-sustaining, and self-regulating system (Isenberg, 2016), in addition to a series of interrelated properties influencing the development of an entrepreneurial ecosystem, including non-linear interactions, environmental adaptability with a certain degree of (in)sensitivity to initial conditions, and the emergence of successful entrepreneurial actors whose governance mechanisms co-evolve within multidimensional geographic boundaries (Han et al., 2021).

Complexity theory becomes particularly relevant also in evolutionary studies on entrepreneurial ecosystems. In this vein, Haarhaus et al. (2020) focus on the complex and nonlinear dynamics of entrepreneurial ecosystems using the framework of chaos theory. Their results demonstrate that chaotic discontinuities are common occurrences during the ecosystem lifecycle. Specifically, phases of heightened chaoticity and complexity are associated with economic crises and the implementation of local policies. Cloutier and Messeghem (2022) instead, characterize the evolution of an entrepreneurial ecosystem as a path dependent process in which a small number of sub-ecosystems coevolve across different sequences. Each sequence is divided into three phases: first, a triggering event (impulse phase); second, the emergence of distinctive attributes (creation phase); third, the establishment of self-reinforcing mechanisms (structuring phase). Throughout the entire process, from the impulse phase to the structuring phase, the entrepreneurial ecosystem exhibits a certain degree of path dependence.

In their study, Crawford et al. (2015) demonstrate that entrepreneurial resources, cognitions, actions, and environments do not adhere to Gaussian distributions, while they follow power-laws typical of complex networks and systems. Such distributions are related to mechanisms affecting entrepreneurship at various levels, including preferential attachment, self-organized criticality, hierarchical modularity, and multiple scale-free causes. For this reason, a network perspective is often complementary to complex adaptive systems theory in the investigation of entrepreneurial ecosystems.

Neumeyer et al. (2019) are among the first to adopt a social network perspective to investigate the complex patterns of social connectivity within entrepreneurial ecosystems. The main peculiarity of their work lies in the proposal of a framework that simultaneously incorporates micro-, meso- and macro-level layers of entrepreneurial ecosystems, enabling the analysis of the interactions among actors across different scales, resembling a multiplex network structure. Similarly, Scott et al.

(2021) represent entrepreneurial ecosystems as multiple interconnected networks. A network-based view of entrepreneurial ecosystems allows authors to capture their vitality, vibrancy, and wealth creation mechanisms arising from complex interacting behaviors. Such networks, in turn, are not static but evolve over time depending on changing states and conditions. Specifically, the transition between two different states is determined by responsive actions to critical situations in the changing environment.

The network structure of entrepreneurial ecosystems makes intermediary organizations a key resource to spread information and knowledge within the network, and to facilitate the access to the venture capital market by entrepreneurs. Business accelerators and universities play a critical role in emerging entrepreneurial ecosystems. In particular, business accelerators forge a broader network of relationships with actors outside of the system, which in turn increase the capacity of the system itself and embed it within a global innovation system (Pustovrh et al., 2020). On the other hand, universities play a strategic role as drivers of regional economic growth by establishing and supporting university spin-off companies (Fuster et al., 2019), and acting as local and regional economic engines with the potential to generate new and disruptive technologies (Chan and Farrington, 2018).

Finally, further aspects are explored by Ives and Carpenter (2007), whose work highlights the interdependence between stability and diversity within ecosystems, and by Jacobides et al. (2018) that introduce an important but neglected characteristic of ecosystems, that is the presence of modularity as a condition that allows at least some degrees of explicit coordination, thus creating the opportunity for an ecosystem to emerge.

Alongside qualitative approaches, other authors analyze the concept of entrepreneurial ecosystem in a more principled way. Isenberg (2011) develops one of the first frameworks attempting to summarize the main elements of an entrepreneurial ecosystem. His model distinguishes six domains: finance, policy, culture, markets, human capital, and supports. Feld (2012) emphasizes the interaction between the players in the ecosystem and the access to all kinds of relevant resources, with an enabling role of the government at the background. In particular, his study shows a list of nine attributes that a successful community of start-ups should have to be an entrepreneurial ecosystem. Such attributes are: leadership, intermediaries, network density, government, talent, support services, engagement, companies, and capital. Foster et al. (2013) determine which pillars of an entrepreneurial ecosystem are the most important for the growth and success of companies. The proposed pillars are: accessible markets, human capital, funding and finance, support systems, regulatory framework and infrastructure, education and training, universities, and cultural support. Stangler and Bell-Masterson (2015) propose four indicators to measure an entrepreneurial ecosystem, that are density, fluidity, connectivity, and diversity. Stam (2015) identifies the framework and systemic conditions of the ecosystem that lead to entrepreneurial activities and to new value creation. The framework is extended in (Stam and Van de Ven, 2021) that introduce an integrative model of entrepreneurial ecosystems consisting of ten elements: formal institutions, culture, network, physical infrastructure, finance, leadership, talent, knowledge, intermediary services, and demand. Roundy (2017) combines entrepreneurship and management research to argue that entrepreneurial ecosystems are influenced by two dominant institutional rules: entrepreneurial market logic and community logic. The author also argues that hybrid support organizations - such as incubators, accelerators, and small business development centers - play a unique role in entrepreneurial ecosystems by exposing participants to the two guiding rules and promoting the presence of a greater diversity of venture types.

1.3 Network-based principles

According to Wurth et al. (2022), there are four main research streams that still need to be investigated in the entrepreneurial ecosystem literature, i.e., context, structure, microfoundations, and complex systems. The authors also include four cross-sectional themes in their proposal for an “Entrepreneurial Ecosystem Research Program”, that are methodologies and measurements, theory, critical research, and transdisciplinary research. In this work, our aim is to quantify structural elements of entrepreneurial ecosystems emerging from the literature by means of metrics from network analysis (Wasserman and Faust, 1994; Humphries and Gurney, 2008; Barabási, 2016; Newman, 2018). Thus, based on the research framework proposed by Wurth et al. (2022), we focus on the cross between structure and methodologies and measurement.

First of all, being our approach innovative to the field, we need to introduce some theoretical prerequisites. Specifically, we combine elements from complexity theory and Social Network Analysis (SNA) to derive the network-based principles that characterize the structure of entrepreneurial ecosystems.

Complexity theory regards the identification and analysis over time of complex systems, including ecosystems, in which the constituent elements give rise to the collective behaviors of the system (Funtowicz and Ravetz, 1994; Thurner et al., 2018). Such complex systems can be described through their structural characteristics (e.g., organizations member features, behaviors, and interaction dynamics) and modeled as networks of interacting entities (Barabási, 2016; Newman, 2018; Russell and Smorodinskaya, 2018). In complex systems, interactions are usually non-uniform and heterogeneous but interactions between elements can be specific (Thurner et al., 2018).

Networks are the preferred tool for mapping such interactions and through different methods attributed to several disciplines, such as mathematics, statistics, physics and computer science (Börner et al., 2007), permit to understand structures, roles and dynamics of complex systems. From this perspective, a network can be considered as an abstraction of observable reality able to explain the performance of real systems since it correlates form with functions, and structure with behaviors (Lewis, 2009; Cerqueti et al., 2018). Hence, network analysis can be useful to describe and analyze the structure and behaviors of several complex systems found in the real world, and to systematically explore performance drivers exploiting concepts such as emergence, adaptability, self-organization, resilience, and flexibility. In this respect, the rapidly increasing mass of data that has become available in many different domains contributes to making the empirical investigation of such complex systems more and more suitable at affordable efforts (Hassanien et al., 2015).

Among the different methods in network analysis, one of the most used tools is SNA. It is an instrument to conceptualize and investigate connections among social entities. In general terms, SNA can be considered as an archetype that abstracts social life in terms of connection structures among entities (Hu et al., 2015) and measures of centrality (Scott and Carrington, 2011). SNA has its foundations in graph theory, a discipline whose birth is attributed to the resolution of the famous "Könisberg Bridge Problem" (West et al., 2001). In general terms, a graph G is described by a couple (V, E), where V is the set of vertices (or nodes), and E is the set of edges (or links). Relationships between any pairs of nodes are described mathematically by a $n-$square matrix A, called adjacency matrix. An element a_{ij} of A is equal to 1 if node i is adjacent (i.e., linked) to node j, and 0 otherwise. The degree of node i is equal to the number of its adjacent nodes, and it is denoted as k_i. Alternatively, let us assume that G is a weighted graph, i.e., its edges are weighted based on factors such as the frequency of occurrence of the associated relations. In this case, G is described by a weighted adjacency matrix W, whose elements are equal to the weight w_{ij} if $(i, j) \in E$, and 0 otherwise. Thus, a network is ultimately represented by a graph, in which nodes correspond to actors and edges stand for specific directed or undirected relationships between them. The use of network analysis to investigate the relationships within an entrepreneurial ecosystem is an innovative field of research (Cavallo et al., 2019). Some recent applications can be found in the literature on entrepreneurial and innovation ecosystems (e.g., Panetti et al. (2020); Cavallo et al. (2021)).

In the end, our process leads to the definition of seven network-based principles characterizing the structure of entrepreneurial ecosystems, which are introduced and described as follows. In more detail, we identify seven distinct elements in the literature corresponding to structural characteristics of entrepreneurial ecosystems. Each element is associated with a network metric based on the meaning of that metric and its subsequent interpretation in an economic environment. These principles are also resumed in Table 1.1, together with the associated network metric and a brief explanation of the latter one; we also report the major contributions on entrepreneurial ecosystems that highlight the relevance of the related element. It is important to specify that the network considered in this work is unweighted and directed; thus, the proposed metrics are appropriate for this type of network. However, the principles can be easily operationalized for weighted and undirected networks as well.

Each network-based principle needs to be interpreted according to a baseline scenario in order to understand if the values we observe are either exceeding or underperforming benchmark values. In network theory, the baseline scenario is usually represented by a set of Erdös Rényi (ER) random graphs. In graph theory indeed, the ER random graph is one of the most commonly used reference models. According to this model, a random network having N nodes is created, and then M edges are generated by pairing among nodes randomly and uniformly (Erdős and Rényi, 1959, 1960). Comparing the obtained results to this benchmark model allows to identify underlying mechanisms in the observed system that depart from the scenario expected by chance (Peel et al., 2017).

1.3.1 Connectivity

Connectivity is an important element for an entrepreneurial ecosystem since the involvement of all the actors is a key factor for a virtuous environment. *Connectivity* affects the possibility to interact between all the members of a system.

We associate *connectivity* to *small-world-ness index*. Recently, attention toward small-world networks has significantly increased due to the peculiarity of their structural properties and their ability to model real systems. Various measures have been proposed to assess this characteristic, starting from the seminal papers by Watts and Strogatz (1998), and Newman and Watts (1999a,b) toward the most comprehensive indicator proposed by Clemente et al. (2018), which computes the small-world-ness of a network by comparing its distance pattern from regularity and randomness to that of a selected small-world model. In this paper, we consider a standard version of the indicator based on the trade-off between a high global clustering coefficient and a short average path length (Humphries and Gurney, 2008). Specifically, it is defined as follows:

$$S = \frac{\gamma}{\lambda} \tag{1.1}$$

where γ is the standardized clustering coefficient, and λ is the standardized average shortest path. Standardization is made by comparing the respective values in the analyzed network with the mean of the corresponding measures over a set of ER random graphs. The clustering coefficient corresponds to the probability that the adjacent nodes of a node are connected, thus capturing the degree to which the neighbors of every node in the network link to each other. The average shortest path is instead the mean of all the shortest paths - that is the path with the fewest number of links - between any couple of nodes. The notion of "small-world" relies on the idea of "six degrees of separation", a popular concept stemming from the experiment conducted by the psychologist Stanley Milgram in the 1960s, according to which the social distance between any two individuals is definitely short (Milgram, 1967). Specifically, a network is considered to display a small-world structure if $S > 1$.

Networks characterized by small-world structure show the presence of hubs connecting a large number of small degree nodes, thus creating short distances between them. For this reason, this metric is appropriate to quantify how well a system is connected considering both cohesiveness and mean distance among nodes.

1.3.2 Density

Density is required as entrepreneurial ecosystems necessitate a high number of interactions among their members to facilitate the establishment of reliable relationships and a frequent knowledge exchange. *Density* determines the frequency of the interactions between the members of a system.

We associate *density* to *mean degree*. The degree of a node corresponds to the number of links incident upon the node. Thus, mean degree represents the average number of connections established by a general node in the network. A distinct behavior of real networks consists in the emergence of giant components. According to Barabási (2016), these dynamics take place in the so called "supercritical regime", i.e., when $\langle k \rangle > 1$, where $\langle k \rangle$ is the mean degree of the network. The network is instead in the "connected regime" if $\langle k \rangle > lnN$, where N is the number of nodes of the network. Therefore, we evaluate the density of the ecosystem by comparing the mean degree to these two thresholds.

In the end, the mean degree reveals if the members of a system frequently exchange information each other.

1.3.3 Stability

Stability refers to the identification of ecosystems as robust structures endowed by the capacity of solving complex and unpredictable problems. *Stability* implies strong connections among the members of a system, making the system able to recover from shocks and to survive over time.

We associate *stability* to *network robustness*. The study of a system's stability via *network robustness* consists in investigating variations in the underlying network topology (Albert et al., 2000; Barabási, 2016; Thurner et al., 2018; Liu et al., 2022). According to Barabási (2016), the breakdown of a network following a random node removal is not gradual. Indeed, removing a limited number of nodes has reduced impact on network topology. However, once the fraction of removed nodes (f) exceeds a critical threshold (f_c), the network becomes suddenly disconnected. Following the Molloy-Reed criteria, the critical threshold is defined as:

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \tag{1.2}$$

where $\langle k \rangle$ and $\langle k^2 \rangle$ are the mean and the second moment of nodes' degree, respectively.

A network shows enhanced robustness if its critical threshold is greater than the randomized network predicition, i.e., $f_c > f_c^{ER}$, where f_c^{ER} is defined as follows:

$$f_c^{ER} = 1 - \frac{1}{\langle k \rangle} \tag{1.3}$$

Then, the stability of a system can be determined by analyzing changes in the dimension of the giant component when random failures occur.

1.3.4 Leadership

Leadership, intended as the presence of a strong group of entrepreneurs and companies, is another fundamental aspect of an entrepreneurial ecosystem. The existence of leaders provides directions and role models, and it is critical in building and maintaining a healthy environment as well as in guiding the community to exploit attractive opportunities.

We associate *leadership* to *page rank centralization*. Page rank is a centrality measure ranking the nodes according to the number and the quality of their connections and the connections of their neighbors (Page et al., 1999), and it is also the fundamental logic behind the rationale of the Google algorithm. Page rank centralization computes the graph-level centrality score, i.e., how much the network is centralized, based on the aforementioned centrality measure. Specifically, the normalized page rank centralization P of a general network G is defined as follows:

$$P(G) = \frac{\sum_i (max_w p_w - p_i)}{\sum_i (p_{max} - p_i)} \tag{1.4}$$

where p_i is the page rank of vertex i, $max_w p_w$ is the maximum value of page rank in the network, and p_{max} is the maximum theoretical page rank for a node in a network with the same number of vertices (Freeman, 1978; Wasserman and Faust, 1994).

Thus, leadership patterns emerge in the presence of a page rank centralization that is higher than the mean of the same metric over a set of ER random graphs.

1.3.5 Diversity

Diversity is defined as the presence of different classes and venture types, whose members interact regardless the category they belong to. *Diversity* contributes to the establishment of a wealthy, heterogeneous, and rich environment, facilitating the cross-fertilization of ideas, and sharing expertise among actors.

We associate *diversity* to *assortativity.* According to Newman (2018), a network is assortative if a significant fraction of links run between nodes of the same type. This metric is defined as the Pearson correlation coefficient of degree or attributes between pairs of linked nodes (Newman, 2002), then ranging between -1 and 1. Thus, assortativity can be calculated based on different characteristics of a node, revealing the extent to which members of different nature (e.g., metadata referring to different domains, such as countries) and different relevance (e.g., in terms of number of connections) interact with each other (Cinelli et al., 2020).

Therefore, a diverse and heterogeneous environment is expected to show neutral behaviors avoiding preferential mechanisms.

1.3.6 Intermediaries

Intermediaries, embodied by support organizations such as accelerators, incubators, universities and research centers, are emphasized as a key resource for an entrepreneurial ecosystem. *Intermediaries* carry out two crucial activities: they act as brokers between different actors of a system, and they channel the information flow toward peripheral members.

We associate *intermediaries* to *betweenness centralization.* Betweenness centrality is a centrality measure determining the number of times a node lies on the shortest path between any two other nodes, thus channeling and controlling the exchange of knowledge in the network (Newman, 2018). Betweenness centralization computes the graph-level centrality score based on the early introduced centrality measure. Specifically, the normalized betweenness centralization B of a general network G is defined as follows:

$$B(G) = \frac{\sum_i (max_w b_w - b_i)}{\sum_i (b_{max} - b_i)} \tag{1.5}$$

where b_i is the betweenness of vertex i, $max_w b_w$ is the maximum value of betweenness in the network, and b_{max} is the maximum theoretical betweenness for a node in a network with the same number of vertices (Freeman, 1978; Wasserman and Faust, 1994).

Then, distinct intermediaries connecting different network components emerge if the value of betweenness centralization is higher than the mean of the same metric over a set of ER random graphs.

1.3.7 Feedback loops

Feedback loops, caused by upward and downward links, are an essential characteristic of an entrepreneurial ecosystem, showing the interdependence between the actors of the system. This notion derives from the original concept of ecosystem, which is generally an environment based on mutual relationships (e.g., the carbon balance of the terrestrial ecosystems).

We associate *feedback loops* to *reciprocity coefficient*. This metric computes the probability that a link from a node to another is reciprocated within the network (Wasserman and Faust, 1994). Reciprocity coefficient determines the fraction of mutual relationships between the members of a system.

Specifically, feedback loops frequently occurs if the reciprocity coefficient is higher than the mean of the same metric over a set of ER random graphs.

For the sake of clarity, we summarize our research framework in Figure 1.1.

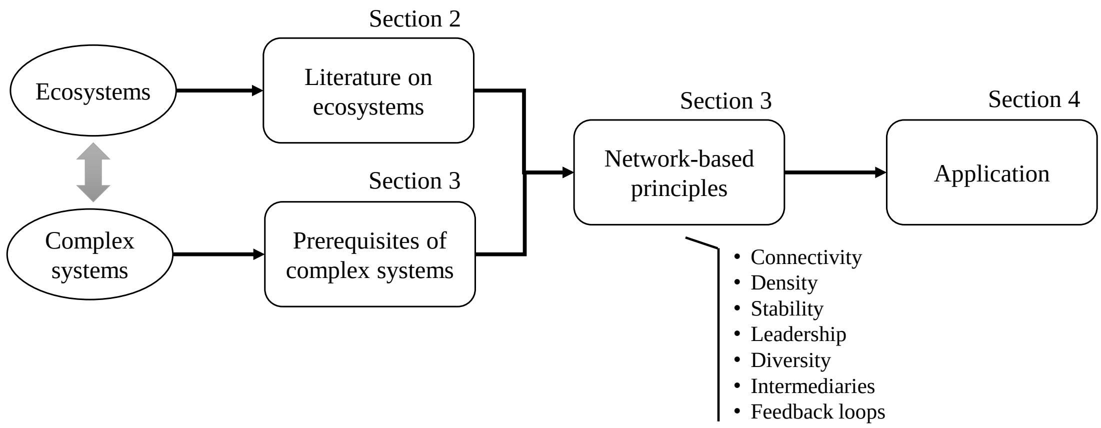


Figure 1.1: Research framework

1.4 Case study: an international start-up network on Twitter

We present an empirical case study about the assessment of the network-based principles of an entrepreneurial ecosystem, taking into account an original network made up of some among the most popular and promising start-ups from seven different countries: Austria, France, Germany, Italy, Spain, Switzerland, and the United Kingdom. In this way, we aim to measure the structural elements of a potential entrepreneurial ecosystem built by using the relationships occurring in April 2020 among a final set of 612 selected start-ups, considering Twitter as a proxy for connectivity.

We consider the analysis of an online network an intriguing case study for many reasons. At first, the pandemic changed the way in which most of the relationships (especially in a working environment) occur (Almeida et al., 2020). Digital platforms have been emerging as essential to

Table 1.1: Network-based principles of entrepreneurial ecosystems

Principle	References	Metric	Definition
Connectivity	Feld (2012), Stangler and Bell-Masterson (2015), Malecki (2018), Stam and Van de Ven (2021).	Small-world-ness index	Relies on the global clustering coefficient of the network and on its average shortest path length and reveals if the network is characterized by a small-world structure.
Density	Feld (2012), Stangler and Bell-Masterson (2015).	Mean degree	Average number of connections of a node in the network. It considers both in- and out-degree.
Stability	Ives and Carpenter (2007), Stanley and Briscoe (2010), Tan et al. (2020).	Network robustness	The network level of tolerance against failures tested by progressively removing random nodes.
Leadership	Feld (2012), Stam (2015), Miles and Morrison (2020), Stam and Van de Ven (2021).	Page rank centralization	Network level of centralization based on page rank. According to page rank, a node is important if it is linked with other important and parsimonious nodes, or if it is highly connected. Page rank takes into consideration direction and weight of a link.
Diversity	Ives and Carpenter (2007), Stangler and Bell-Masterson (2015), Roundy (2017).	Assortativity	Determines the tendency to connect between similar nodes based on their degree and attributes. This coefficient ranges between -1 - completely disassortative - and 1 - completely assortative.
Intermediaries	Feld (2012), Roundy (2017), Fuster et al. (2019), Pustovrh et al. (2020), Stam and Van de Ven (2021).	Betweenness centralization	Network level of centralization based on betweenness centrality. This centrality measure identifies those nodes which lie on many shortest paths.
Feedback loops	Heimann and Reichstein (2008), Stam (2015).	Reciprocity coefficient	Proportion of mutual links in a directed graph. It highilghts the existence of reciprocal connections between two members of the network.

communicate, and they will probably keep this key role in the future as well (Seetharaman, 2020). As a consequence, the concept of proximity, which is essential in the definition of entrepreneurial ecosystems, has changed. Rather than geographical and physical, we need to consider other kinds of proximity, and our case study provides an example of that. Secondly, Twitter is one of the most used social networking platforms among innovative firms, and its usage increases networking capabilities as well as the possibility of establishing professional links between different companies. Overall, the use of social media has a positive and significant impact on an organization's performance by reducing costs, enhancing customer relations, and improving information accessibility (Tajudeen et al., 2018). SNA is a powerful tool to investigate the entrepreneurial success of start-ups. In particular, the social proximity to influential actors is an efficient proxy to predict the success of start-ups, with the quality of neighbors being more significant than their quantity (Gloor et al., 2013). The analysis of following and followers structure on Twitter especially, provides insights into the the overall company performance. Indeed, the more central the position of a start-up in the online community, the higher its probability of success (Perotti and Yu, 2015). Some recent works related to business are exploiting information derived from Twitter in order to gain relevant insights about companies. For instance, Tumasjan et al. (2021) study the role of digital traces on Twitter as a predictor of business venturing opportunities, finding that Twitter is particularly important when start-ups are still young. Yue et al. (2019) inspect the engagement among Fortune 200 and top startup CEOs on Twitter. A more comprehensive review of the importance of social media for entrepreneurship can be found in (Olanrewaju et al., 2020). Finally, analyzing relationships occurring in an online platform allows us to go beyond regional and national borders - as entrepreneurial ecosystems are usually defined - and deal with an entrepreneurial ecosystem at the continental level.

In the Twitter network, a following relationship can be considered as an expression of interest toward another actor in the sense that it expresses the desire to be up-to-date with the news posted by the followed actor. Indeed, the following relationships on Twitter are among the main factors that rule the timeline, i.e., what the users see on the platform, and, in turn, the information they automatically have access to.

In order to define formally the relationships on Twitter, let us consider a generic couple of start-ups, A and B. If A is a follower of B it means that A has an interest in B. Similarly, if B is a follower of A it means that B has an interest in A. If A and B follow each other then A is accounted in the list of B's followers while B is accounted in the list of A's followings. Accordingly, being the relationship of interest mutual, B is accounted in the list of A's followers while A is accounted in the list of B's followings.

Data processing, network analysis and all simulations are conducted using the software R (R Core Team, 2014), specifically the "igraph" package (Csardi and Nepusz, 2006). Network extraction is realized by interacting with Twitter's API through the "rtweet" package (Kearney, 2019).

1.4.1 Data collection

The process of data collection follows a three-steps procedure.

Step 1

The lists of the most promising start-ups for each country in 2019/2020 are collected from accredited sources which provide every year the names of the most noteworthy start-ups in the international context. In particular, these data sources indicate fast-growing start-ups according to various criteria, including funding raised, revenues generated, capital invested, number of employees, industry, and business model. Among these start-ups, we select those with a website up and running at the time of data collection. The list of the sources used within this study is the following: Sifted, Seedtable, StartupItalia, Top 100 Swiss Startup Award, Startups, StartupsReal by El Referente. A description of each source as well as the related access date are provided in Appendix A.1.

Step 2

All the start-ups identified during the previous step are classified in fifteen different categories according to their business activities. The identification criteria are based on the description of each category, which is reported in Appendix A.1.

This classification is obtained in a two-fold way: some sources from which start-ups are taken explicitly indicate the category of each start-up; for the remaining cases, the category is retrieved from the description of the company as provided either on its website or on its Twitter profile. It is important to specify that we consider the main business sector for each start-up, thus avoiding the attribution of multiple categories to the same node for computational reasons.

Step 3

We exclude all the start-ups that do not own a Twitter account, as being Twitter the proxy we refer to build the network. The final dataset comprises 612 start-ups out of an initial set of 712 start-ups. The distribution of the selected start-ups over different countries and categories is represented in Figure 1.2 and Figure 1.3, respectively.

1.4.2 Network construction and visualization

The network of start-ups is built by downloading follower and following relationships for each of the 612 accounts, through the interaction with Twitter's Standard API.

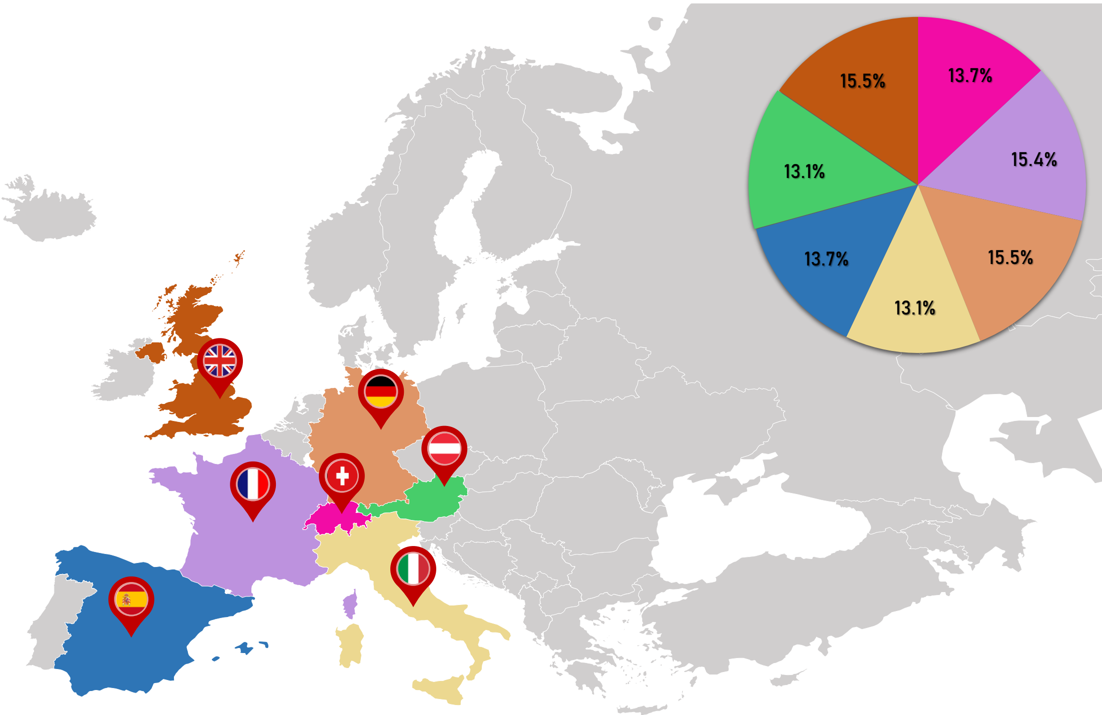


Figure 1.2: Distribution of start-ups over countries

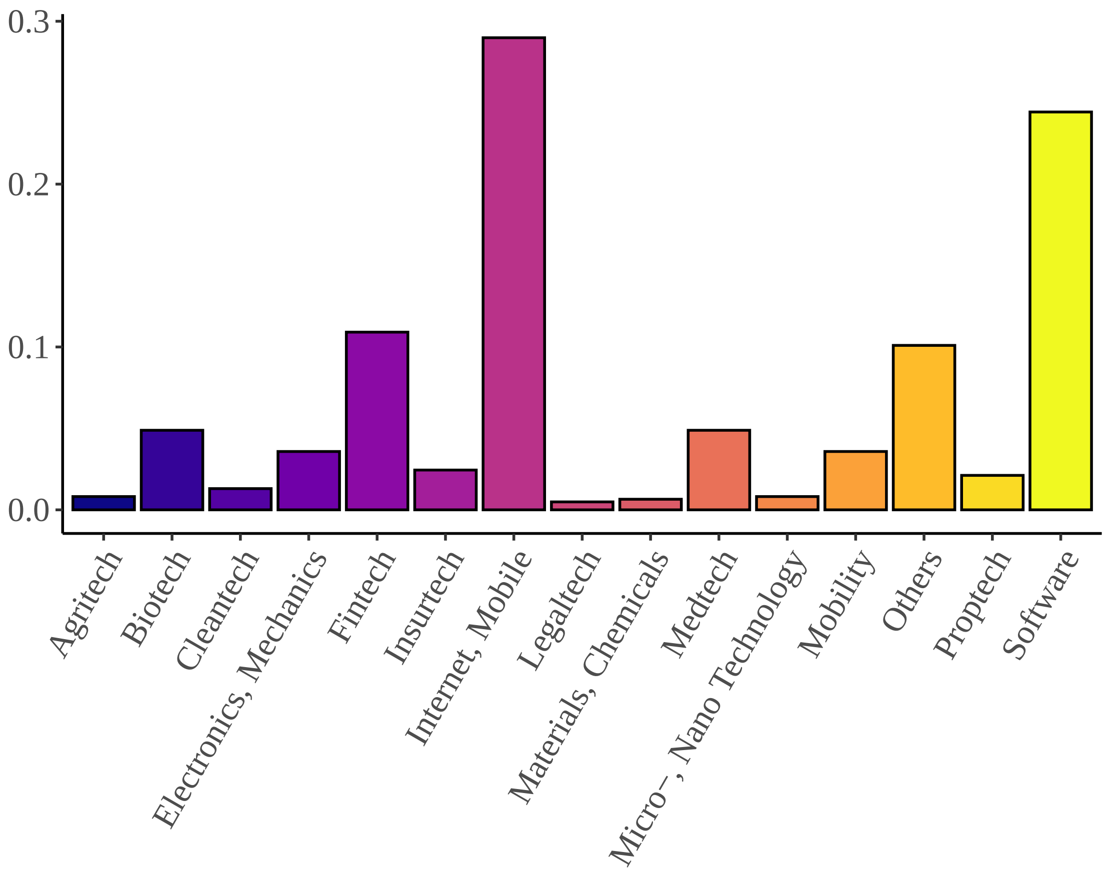


Figure 1.3: Distribution of start-ups over categories

The resulting network includes 612 nodes, corresponding to the aforementioned set of start-ups, and 1,150 directed edges. In more detail, each start-up corresponds to a node, while an edge stands for the relationship between two different start-ups, i.e., there is a directed link from i to j if *start-up* i follows *start-up* j on Twitter, while there is a directed link from j to i if *start-up* j follows *start-up* i (or, equivalently, *start-up* i is followed by *start-up* j) on Twitter. The bulk of these links take place among 419 start-ups, representing the giant component of the network, i.e., the greatest subset of nodes all mutually connected by any path. More precisely, the network is structured in the following way: a giant component, made up of 419 nodes; eight small connected components, six of which are made up of 2 nodes, while the other ones are made up of 3 and 5 nodes; the remaining components consist in 173 isolated nodes. The network is displayed in Figure 1.4.

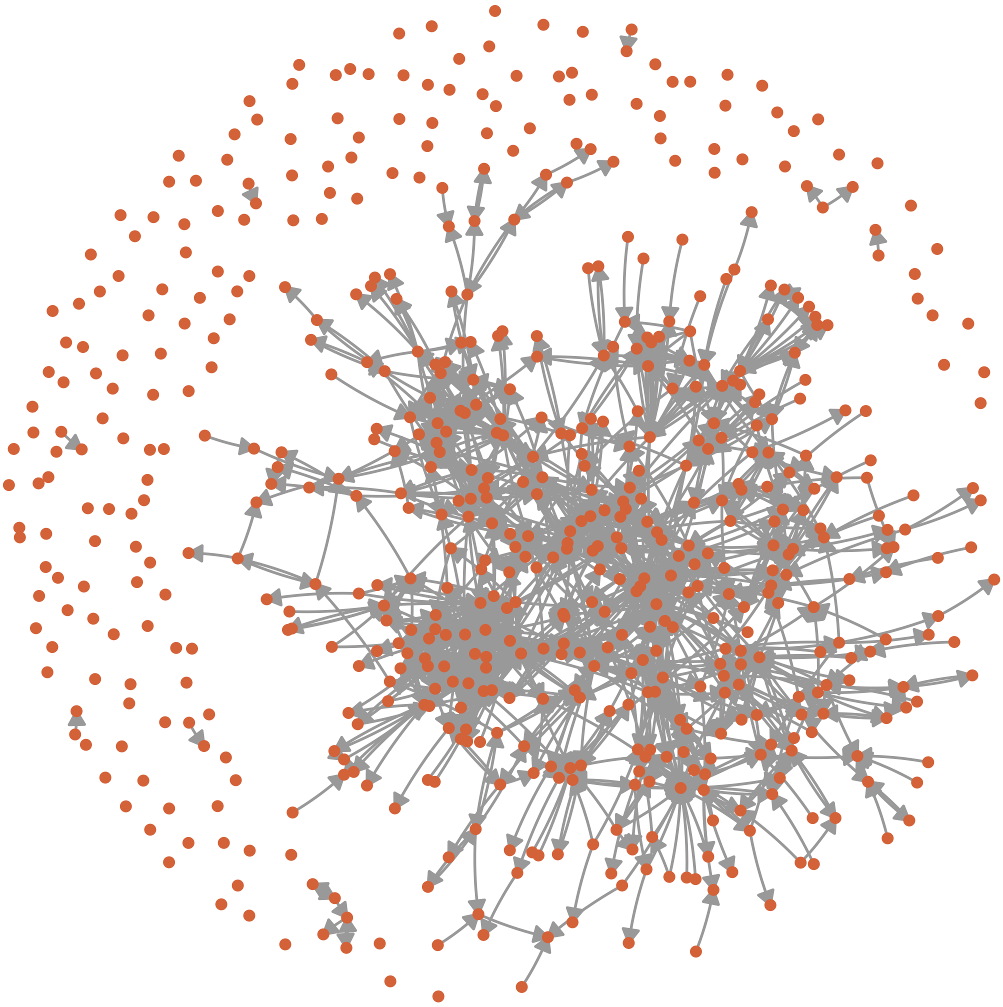


Figure 1.4: Network of European start-ups on Twitter

1.5 Empirical results

The following section shows the results we obtain by measuring the seven network-based principles of the network introduced in Section 1.4. Hereafter, we address each principle formulated in Section 1.3, and then we summarize all the results in Table 1.2.

Connectivity

We standardize clustering coefficient and average shortest path by comparing them with the mean of the corresponding values over $1,000$ ER random graphs characterized by the same number of nodes and edges. The computed small-world-ness index results in 33.7, thus suggesting the network to be properly small-world.

Density

The mean degree of the network is equal to 3.76. According to Barabási (2016), since the mean degree is greater than 1, the network is in the "supercritical regime". However, the obtained value is lower than $lnN = 6.42$, i.e., the critical threshold identifying the "connected regime".

Stability

According to the Molloy-Reed criteria, 86% of the actual nodes need to be removed from the network of start-ups to disconnect it, based on its actual topological structure (i.e., $f_c = 0.86$). This is greater than the threshold predicted for the randomized network, which is 0.73. The enhanced robustness of the network emerges also from Figure 1.5, where we represent the portion of nodes belonging to the giant component ($P_\infty(f)$) while progressively removing an increasing fraction of random chosen nodes (f), with respect to the initial portion of nodes in the network belonging to the giant component ($P_\infty(0)$). Since the error tolerance analysis could be influenced by the nodes that are randomly chosen at each step of the algorithm, this kind of simulation is repeated $1,000$ times, in order to smooth out eventual fluctuations of the outcome. The evaluation of network robustness is then enriched by the comparison with the average results obtained over $1,000$ ER random graphs with the same number of nodes and edges.

The results are strongly influenced by the structure of the network, since the huge amount of small degree nodes allows the network to efficiently react in case of random failures. Therefore, the connections among the start-ups are stable enough to guarantee that the size of the giant component will not considerably decrease, and that the network will not become highly disconnected.

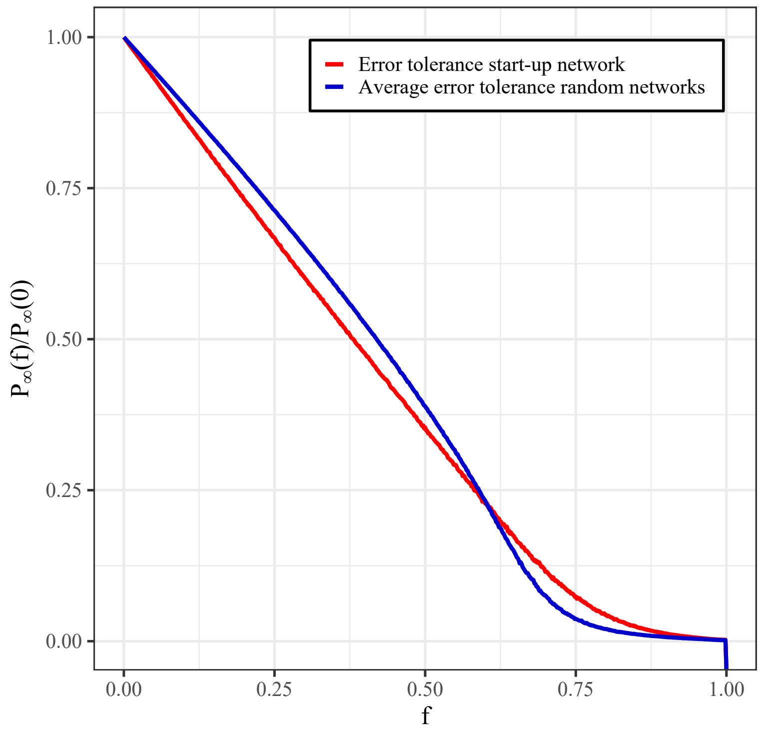


Figure 1.5: Error tolerance analysis: effects on the giant component. $P_\infty(f)$ is the number of nodes belonging to the giant component over the total amount of nodes in the network, when removing a specific fraction of nodes f; $P_\infty(0)$ is the number of nodes belonging to the giant component over the total amount of nodes in the network, when removing 0 nodes, i.e., the initial portion of nodes in the network belonging to the giant component. $P_\infty(f)/P_\infty(0)$ represents the portion of nodes belonging to the giant component when removing a certain fraction of nodes f, normalized by the initial portion of nodes belonging to the giant component.

Leadership

The network level of centralization based on the values of page rank is equal to 0.02. The average value of centralization over 1,000 ER random graphs characterized by the same number of nodes and edges is 0.006. Therefore, the network structure shows the presence of leadership patterns.

Diversity

Concerning the assortativity coefficient to degree, the network results to be quite neutral, being the coefficient equal to -6.8×10^{-3}. On the other hand, the values of assortativity to the two attributes of the start-ups, which are country of origin and category, are 0.73 and 0.23, respectively. Therefore, we can observe a clear evidence of the effect of geographical proximity in the establishment of relationships within the network. A quite similar phenomenon is determined by the membership of the start-ups to the same category. On the other hand, the activity of the start-ups in the network does not show preferential mechanisms between similar nodes in terms of their size.

Intermediaries

The network level of centralization based on the values of betweenness centrality is equal to 0.04. The average value of centralization over $1,000$ ER random graphs characterized by the same number of nodes and edges 0.06. Therefore, the network structure does not highlight the presence of distinct intermediaries.

Feedback loops

The reciprocity coefficient in the analyzed network is equal to 0.32, namely about 1/3 of relationships are mutual. This value definitely overperforms the average reciprocity coefficient over $1,000$ ER random graphs that is 3.1×10^{-3}. Then, we can conclude that the presence of feedback loops is frequent within the network of start-ups.

Table 1.2: Results obtained by measuring the seven network-based principles in the network of start-ups, compared to benchmark values.

Metric	Result	Benchmark values	Interpretation
Small-world-ness index	33.7	1	The network shows a small-world structure.
Mean degree	3.76	1 (supercritical) 6.42 (connected)	The network depart from the ER random scenario, but it is not highly-connected.
Network robustness	0.86	0.73	The network displays enhanced robustness.
Page rank centralization	0.02	0.006	The network structure reveals the presence of leadership patterns.
Assortativity	−0.007 (degree) 0.73 (country) 0.23 (category)	−1 (disassortative) 0 (neutral) 1 (assortative)	Clear evidence of geographical proximity, and sectorial clusterization, while no preferential mechanisms between nodes of the same size.
Betweenness centralization	0.04	0.06	The network structure does not highlight the presence of distinct intermediaries.
Reciprocity coefficient	0.32	0.003	Structural feedback loops occur frequently in the network.

1.6 Conclusions and discussion

In this paper, we propose an innovative approach to the measurement of the structure of entrepreneurial ecosystems. Since entrepreneurial ecosystems can be considered complex systems characterized by a multitude of actors frequently interacting among them, we adopt a network-based approach to analyze the connections between the members of an entrepreneurial ecosystem. Specifically, we address some important structural features that, according to the literature, characterize entrepreneurial ecosystems: connectivity, density, stability, leadership, diversity, intermediaries, and feedback loops.

We associate each element with a consolidated network metric. In doing so, we provide seven network-based principles characterizing entrepreneurial ecosystems that can be measured and assessed to monitor and evaluate the structure of a given entrepreneurial ecosystem. In this way, we shed light on an open issue in the current literature, which is the investigation of the structure of entrepreneurial ecosystems. In particular, the measurement of the structural features of an entrepreneurial ecosystem is of fundamental importance to tailor proper policies fostering collaboration and innovation, since all the actors, structures, and practices that characterize entrepreneurial ecosystems strongly contribute to digital innovation and economic growth.

This is a novel approach to the measurement of entrepreneurial ecosystems with respect to previous methods. For instance, considering the EIDES family of indicators, there are two main novelties introduced by our approach. Firstly, the indicators characterizing the EIDES correspond to statistics published (in almost all cases) annually by international institutions, such as Eurostat, World Economic Forum, and others. It follows that, the EIDES represents a well-recognized ex-post indicator aimed at measuring the digital development of entrepreneurial ecosystems. Differently, the proposed approach, being based on network metrics, allows monitoring the structure of an entrepreneurial ecosystem in real-time, without the need to wait for the publication of national and international statistics. Secondly, as linked to country-level statistics, the EIDES does not allow dealing with entrepreneurial ecosystems at the international level. Even if entrepreneurial ecosystems are usually defined in terms of local districts at the regional or national level, the investigation of the establishment of innovation and digital ecosystems at the international level is gaining attention recently, especially in a European policy context. Our approach, being based on connections between organizations, can be applied in an international context in all the cases in which the analyzed ecosystem is characterized by interactions between foreign partners, as in the case study we report. Nevertheless, in the future we intend to create an overall index including all the network-based measures, being coherent with the EIDES family of indicators.

We also contribute to the literature by proposing an alternative conceptualization of ecosystems considering online relationships. In particular, we provide an example of a network constituted by 612 start-ups belonging to seven different countries, built by considering the relationships occurring among them on Twitter. This is an innovative way to investigate entrepreneurial ecosystems that can gain increasing attention in the next future. The pandemic indeed, shifted several relationships toward digital platforms, extending the concept of proximity that is central in the definition of entrepreneurial ecosystems. In particular, Twitter is one of the most important social networking platforms among high-technology firms, and it has been recently used as a predictor of start-ups' success (Tumasjan et al., 2021). Finally, online relationships can extend beyond regional and national borders, allowing the establishment of entrepreneurial ecosystems at the continental level.

As a further development, we would like to integrate the network of start-ups on Twitter with other kinds of network including other types of actors to investigate their interactions in a multilayer ecosystem. In this way, we aim to detect the interrelated dynamics and mechanisms that affect entrepreneurial ecosystems online and offline, simultaneously.

The proposed approach also leaves room for some policy implications. Indeed, it can provide real-time indications on how an entrepreneurial ecosystem is structurally developing, allowing for timely interventions in case of unmet properties. In this way, we contribute to a claimed gap in the current analysis, that is the need of data-driven approaches to model and measure entrepreneurial ecosystems. More generally, the proposed methodology aims to support policy-makers and institutions funding the establishment of collaborative networks and entrepreneurial and innovation ecosystems. The results obtained from the measurement of the network-based principles in fact, can help these actors evaluate the impact of their policies on the structure of a specific system.

Looking at the results, we find that the members of the analyzed network are able to communicate with each other in few steps due to the small-world structure of the system, despite the low number of edges in the network. The average number of connections established by a start-up suggests the network to be in the "supercritical regime", thus departing from the ER random scenario. However, the mean degree is not sufficiently high to allow the network to shift toward the "connected regime". One possible reason of such behavior could be the stage of development of the analyzed network, as well as the early adoption of Twitter in the business processes. In conclusion, our findings highlight a partial ecosystem-like nature of the analyzed network, pointing out those elements that need to be better established in order to consider it a proper entrepreneurial ecosystem. An eventual limitation of our case study is related to the fact that start-ups might be connected via their founders or CEOs on Twitter rather than via their corporate accounts.

This work represents a first attempt to analyze the social network structure of a system to guide policy-making decisions. Although we acknowledge that the obtained results depend on the selected metrics, we consider the ones we test the most efficient in quantifying the related structural properties. For instance, concerning connectivity, we decide to use the small-world-ness index since it is a summary measure taking into account both cohesiveness and mean distance among nodes, thus providing more information about overall connectivity than other network measures, such as connectedness. Additionally, we do not associate density with edge density, as it is a relative measure dependent on the size of the network and does not fully capture the frequency of interactions. Moreover, considering mean degree enables us to conduct the analysis on the entire network, rather than restricting it to the largest connected component. Finally, it is important to specify that in the presence of undirected networks (i.e., links between nodes are not oriented), the existence of leadership patterns can be assessed through eigenvector centralization (Bonacich, 1991), while feedback loops can be identified in correspondence with closed triangles, thus by computing the clustering coefficient of the network. In fact, both page rank centralization and reciprocity coefficients are related to directed networks. In the future we aim to complement the current analysis with community detection methods to identify specific groups of start-ups particularly clustered, and observe whether their composition varies over time. Furthermore, we aim to apply the methodology to other ecosystems in order to collect data for a comparative analysis among different systems.

Chapter 2

How do partner selection strategies affect the amount of funds in collaborative research projects? Evidence using the dual-projection approach

Abstract

The choice of partners is a key strategic lever for the performance of collaborative projects. In this study, we contribute to the existing body of collaborative R&D literature by investigating the effect of different partner selection strategies in collaborative research projects. Specifically, there is a lack of evidence on the impact of collaborating with new over existing partners on the amount of funding received. To address this gap, we adopt an innovative technique in Social Network Analysis, called the "dual-projection" approach. In particular, we analyze the network structure of projects funded by Horizon 2020, the eighth European Framework Programme. Our results indicate that partnering with new organizations increases the probability of obtaining more funds when entering a new project compared to collaborating with already existing partners. Additionally, projects coordinated by private or public organizations are more likely to secure higher funding compared to research centers and higher education institutions. However, the relevance of partners' connections decreases as we move away from the focal organization. Finally, our findings provide significant insights into the effectiveness of the European Research and Technological Development Policy.

Keywords: Collaborative R&D, Partner selection, Research projects, Social Network Analysis, Horizon 2020.

2.1 Introduction

Open innovation has shifted several paradigms in scientific theories and practices, rapidly becoming a valuable concept for many firms and in several different contexts (Huizingh, 2011). According to Chesbrough (2003), the "father" of the open innovation concept, the earlier logic of the closed innovation approach was based on investments in internal Research and Development (R&D), which led to many breakthrough discoveries. On the contrary, one of the main principles of open innovation is that external R&D can create a significant value as well. As a consequence, R&D is increasingly opening up (Enkel et al., 2009), and new ways to support collaboration are emerging (Antikainen et al., 2010).

Meanwhile, the concept of collaborative R&D has become relevant in recent scientific literature. Several works focus on the impact of collaborations on innovation (Hoang and Rothaermel, 2005, 2010), and on the drivers of university-industry collaboration (Bhullar et al., 2019; Atta-Owusu et al., 2021), while other authors are more interested in investigating the dynamics of collaborative relationships (Tatarynowicz et al., 2016; Jakobsen et al., 2019). In this work, we aim to extend the strand of the literature exploring the effects of partner selection strategies in collaborative research projects. In particular, we want to investigate which configuration of participants' connections is more likely to provide collaborative research projects with a higher amount of funds, contributing to deepening the research on the link between collaboration and research funds (Defazio et al., 2009; Ebadi and Schiffauerova, 2015; Ma et al., 2015). In fact, little attention has been paid to the investigation of partner selection strategies in the collaboration literature, whereas there is much more evidence in the field of alliances. Our aim is to fill this gap and shed light on specific aspects in the collaboration literature, i.e., the choice between new and old partners in the field of R&D alliances (Li et al., 2008; Kang and Zaheer, 2018), and the role of social capital in inter-organizational settings (Seo, 2020). It is important to notice that, in the context of collaborative projects, we consider as new partners those organizations that have never participated in a joint project at the time of the analysis, while old partners correspond to organizations that have already collaborated in previous research projects at that time.

To this aim, we adopt an innovative approach in this field. First, we exploit Social Network Analysis (SNA) techniques to analyze the connection structure of the network representing collaborative research projects, where two projects are connected if there exists at least one organization taking part in both of them. SNA is one of the most used methods to conceptualize and investigate connections among social entities (Wasserman and Faust, 1994), and the collaborations between organizations participating in research projects can be easily represented through a network. More specifically, we model the connections between projects through an undirected weighted network, where nodes are projects and links are weighted according to the number of common participant organizations. Second, the main methodological novelty of this study is the adoption of the so called "dual-projection" approach, introduced by Everett and Borgatti (2013) to avoid structural data losses when dealing with two-mode data, as in the case of collaborative projects. This kind of data is characterized by the existence of two different sets of nodes, and only inter-group connections (i.e., there are no intra-group links). In the case of collaborative projects, the two distinct sets of nodes correspond to projects and organizations, respectively. Applying the dual-projection approach to the analysis of collaborative research projects allows us to include the centrality of organizations (i.e., their importance in terms of connections) in the computation of the centrality of projects. This procedure enables to express the dependence of the relevance of projects on the importance of the proposing institutions. In this way, we can compare the results obtained through the application of the dual approach to the results deriving from the so called "classical" approach, which does not account simultaneously for both the centrality of projects and the centrality of participants. The comparison between the two approaches will allow us to investigate different partner selection strategies and understand which one is more likely to lead to a higher amount of funds. A full description of the two methods, as well as their interpretation in the context of partner selection strategies, are reported in the paper.

We focus on a highly relevant case study in the field of collaborative R&D, namely the projects funded by the eighth European Framework Programme (EU FP), Horizon 2020 (H2020). The EU FPs are multi-annual programmes providing funds mainly to EU member states, but also to associate countries, in order to promote long term investments in several areas, in accordance with the establishment of a systematic Research and Technological Development (RTD) policy. Almost all works in the literature using networks to investigate EU FPs focus on the connections between participant organizations (Enger, 2018; Balland et al., 2019; Cinelli et al., 2022), disregarding the connection structure of projects. As introduced above, the adoption of the dual-projection method will allow us to go beyond this approach. Moreover, many policy insights into the mechanisms affecting European research projects have been provided in previous studies, but there is a lack of evidence on the strategic decisions of participant organizations. Thus, this work contributes to the current strand of literature on EU FPs by shedding light on the effects of different partner selection strategies on the likelihood of securing increased funding for H2020-funded projects. In doing so, this study presents relevant managerial implications for participant organizations involved in collaborative research projects, particularly in the field of EU FPs.

The paper is organized as follows. Section 2.2 presents an overview of the state of the art on the phenomenon of partner selection and the EU FPs. Section 2.3 defines the research methodology, stemming from SNA and graph theory, and gives details about the dual-projection approach. The case study and the data collected are introduced in Section 2.4. Section 2.5 describes the econometric model, while Section 2.6 reports the results of the analysis. Finally, Section 2.7 discusses the main theoretical and practical contributions of this paper, as well as the limitations and future aspects we would like to investigate more.

2.2 State of the art and hypotheses

As introduced in Section 2.1, this paper aims to contribute to two different strands of literature. First, we want to provide insights into the effectiveness of different partner selection strategies in the field of collaboration studies. To this aim, we formulate three hypotheses, which are mostly derived and adapted from the analysis of the phenomenon of partner selection in the alliance literature. Second, we deepen the current analysis on research projects funded by the EU FPs, assessing one specific hypothesis for this case study. Thus, in what follows, we introduce the state of the art of both research streams.

2.2.1 Partner selection mechanisms

The phenomenon of partner selection is one of the most investigated in the literature on collaborative R&D and innovation management. There are several determinants behind the choice of partners that need to be distinguished from the motivation to enter into a joint venture (Al-Khalifa and Eggert Peterson, 1999). A non-exhaustive list of partner selection criteria includes financial and economic stability, business know-how, knowledge of the local market, a clear view of the strategic direction of the business, and the existence of chemistry among partners (Doherty, 2009). Moreover, firms' size and research capabilities represent other two factors driving partner selection mechanisms (Mindruta et al., 2016). According to Diestre and Rajagopalan (2012), technological relatedness and development experience significantly increase the likelihood of partnership formation, while trust and commitment are found to be crucial in the presence of high process manageability and outcome interpretability (Shah and Swaminathan, 2008). Ryu et al. (2020) complement the research stream on alliance formation by shedding light on the effect of market overlap and rivalry on collaboration.

The structure of strategic alliances is also relevant. For instance, Vătămănescu et al. (2020) examine the impact of the strategic network structure on the innovation performance of small and medium-sized enterprises (SMEs), highlighting the positive effect of competitive knowledge sharing. Wathne et al. (2018) analyze the problem from a different perspective, focusing on a specific strategic partnership, i.e., between supplier and reseller. They find that stringent reseller selection efforts contribute to promoting incremental supplier investments and protecting them ex-post.

Some authors investigate partner selection and interactions through network analysis. For instance, Weng et al. (2014) explore how technological positions emerge from the interactions among partners, looking at the alliance network structure, while Kang and Zaheer (2018) conceptualize alliance partner choice in terms of network distance. Kumar and Zaheer (2019) instead, analyze the ego network of the focal firm to determine the effect of stability on innovation in a collaborative environment.

Hypothesis 1 *The network structure of connections plays a crucial role in collaborative projects.*

Based on the study of Li et al. (2008), the more ambitious an alliance's innovation goal, the more desirable that partners are friends (i.e., they have previous alliance experience) rather than strangers, while strangers are preferred to acquaintances. In this vein, Kang and Zaheer (2018) analyze different alliance partner choices based on network distance, finding that distant partners offer benefits because of diverse knowledge, existing partners provide advantages in terms of trust and established communication routines, while close partners may not benefit the firm as much as the other categories.

Hypothesis 2 *Establishing connections with existing partners in new collaborative projects provides more advantages compared to partnering with previously unrelated actors.*

Firms with partner selection capability generally achieve positive organizational outcomes in terms of alliance structure and portfolio diversity (Sarkar et al., 2009). Then, the choice of a partner becomes a proper managerial control task, being a condition for the success of a business network (Moeller, 2010). This aspect is supported by Wang and Rajagopalan (2015), which sustain that heterogeneity in alliance outcomes arises because of the different capabilities needed to address challenges and opportunities in the partnership's pre- and post-formation phases. In addition to capabilities, complementary knowledge is one of the most crucial characteristics for prospective partners as well as a key factor for innovation performance (Wu et al., 2009; Baum et al., 2010). To this aim, the combination of inter-organizational learning and social capital are considered strategic elements for a consortium performance (Seo, 2020). Wu et al. (2019) find team diversity to be positively associated with project performance, arguing on the mediating effect of conflicts on the relationship between the two phenomena. The importance of a high level of diversity is also claimed by Gattringer et al. (2017) in relation to collaborative foresight projects. In particular, they highlight that heterogeneity in terms of technological resources is more critical than organizational proximity.

Hypothesis 3 *A high level of diversity among partners is a positive factor in collaborative projects.*

According to Bengtsson et al. (2015), involving a large number of partners increases the level of complexity of the alliance, so managers and researchers need to put more effort into organizational problems rather than into innovation activities. This issue can be mitigated by relying on project management techniques that help to select the right partners, e.g., in an open innovation project (Guertler and Sick, 2021). Moreover, when partners are at the early stage of product development and have limited experience in forming partnerships, as well as different technological portfolios, venture capitalists (VCs) can be intermediaries facilitating R&D collaborations (Reuer and Devarakonda, 2017).

2.2.2 The European Framework Programmes

One of the most interesting case studies in the field of collaborative R&D at the European level is represented by the projects funded by the EU FPs. SNA is one of the most used tools to analyze structures, roles, and dynamics of collaborative networks, and this is also the case of many empirical studies on EU FPs. Roediger-Schluga and Barber (2008) are among the first who investigate the global structural properties characterizing the collaborative networks of the first five EU FPs. Enger (2018) addresses the position of organizations in projects funded by the EU FPs distinguishing between different levels of centrality. In particular, the author finds that higher education institutions (HEIs) with high levels of centrality in the seventh FP (FP7) show a greater propensity of obtaining funds in H2020 compared to HEIs that were not participating in FP7. A more comprehensive use of SNA is exploited by Balland et al. (2019), which investigate the structure of the last three EU FPs by computing global network properties (e.g., global clustering coefficient, assortativity, and inequality coefficient) and the different roles of participants through the use of centrality measures. Their results highlight that participants from EU-15 countries tend to have a higher centrality than participants from EU-13, associated, and third countries in H2020. Moreover, the connection structure also depends on the activity type of the organizations, as HEIs show significantly higher centrality measures than e.g., private companies. Analyzing the Agri-Food network funded by the FPs between 2008 and 2014, de Arroyabe et al. (2021) highlight the importance of heterogeneity and geographic diversity of participants in determining the effectiveness of such innovation systems. To complement this strand of literature, Cinelli et al. (2022) introduce a new metric, i.e., the collective network effect (CNE) to measure the benefit of network membership among the participants in projects funded by the seventh EU FP. In particular, they find that organizations with a higher CNE generally have access to more funds than those with a lower CNE.

Hypothesis 4 *The centrality of European-funded projects in the collaborative research network is positively related with the amount of funds.*

Several authors analyze projects funded by the EU FPs, addressing how spatial characteristics have an impact on the structure of collaborative networks. In particular, Scherngell and Barber

(2011) find that spatial factors significantly affect industrial collaboration, while in the public sector the effect is smaller. Scherngell and Lata (2013) confirm that geographical distance exerts a negative effect on collaboration probability, although the effect significantly decreases in the time span between the fifth and the sixth FPs. These preferential mechanisms make Europe remaining a collection of national innovation systems, far from the realization of the European Research Area (Chessa et al., 2013). Besides geographical factors, economic, technological and social distance affect the probability of two regions to link to each other (Amoroso et al., 2018).

Another emerging aspect is the effect of previous participations in assuming a leadership position in subsequent FPs (Breschi and Cusmano, 2006; Enger and Castellacci, 2016), while previous co-publication activity has only a marginal effect on the probability of being funded (Hoekman et al., 2013). An interesting insight is provided by the study of Cavallaro and Lepori (2021), where authors demonstrate that institutional barriers, such as the Brexit for the United Kingdom and the Switzerland's reclassification as third country, have a negative impact on consortium building mechanisms.

Other works investigate the importance of reputation of applicant organizations in the collaborative networks emerging from the EU FPs. The study of Lepori et al. (2015) analyzes the participation of HEIs in projects funded by the EU FPs, and their association with HEI characteristics. Their results suggest the close relation between HEI reputation and their network structure. Enger and Castellacci (2016) present a timely analysis of participations in projects funded by Horizon 2020, finding that the probability of being funded is enhanced by the scientific reputation of the applicants. Thus, reputation has a direct impact in fostering the establishment of connections in a collaborative environment. This is the reason why institutions with a higher reputation tend to become core actors in the EU FPS collaborative research networks (Breschi and Cusmano, 2006; Heller-Schuh et al., 2011). Calignano (2021) focuses on the reverse effect, finding that cooperation in this kind of inter-organizational networks increases the scientific and academic reputation of organizations. Finally, there are other factors that significantly affect collaboration choices in the EU FPs networks, such as prior acquaintance, and different kinds of proximity, i.e., thematic, geographical, social, and organizational (Paier and Scherngell, 2011; Heringa et al., 2016).

2.3 Research methodology

It is worth recalling that our focus is to investigate a network of collaborative research projects with particular attention to the relation between the position of projects in the network and the amount of received funds. This approach draws elements from SNA in order to proxy both the quality and the quantity of connections of each project in a collaborative network. From this perspective, two projects are connected if at least one organization participate in both projects.

In general terms, SNA can be considered as an archetype that maps social relationships in terms of connection structure among entities (Hu et al., 2015) and measures of centrality (Scott and Carrington, 2011). Social interactions are typically heterogeneous and non-linear, proper of a complex system (Thurner et al., 2018). Complex systems are commonly modeled as networks of interacting entities (Barabási, 2016; Newman, 2018; Russell and Smorodinskaya, 2018), since networks are able to explain the performance of real systems, correlating form with functions, and structure with behaviors (Lewis, 2009; Cerqueti et al., 2018).

SNA has its methodological foundation in network theory. Thus, before going into more detail, we need to recall some definitions from network theory. Moreover, we will discuss the "dual-projection" method, which is one of the main novelties of our approach.

2.3.1 Preliminaries on network theory

A network $N = (V, E)$ is composed by a set V of n vertices, or nodes, and a set E of m edges, or links. As stated before, a network is a representation of some kind of interaction or relationship between entities (e.g., individuals, organizations, or countries). Based on the type of the relationship, the network can be either directed or undirected. Since collaborations between entities are usually represented by undirected networks, in what follows we recall the theoretical foundations for this type of networks. In particular, in an undirected network, it holds that if $(i, j) \in E$, then $(j, i) \in E$. In this case, i and j are said to be adjacent. The degree d_i of the vertex i is the number of its adjacent nodes, and we denote the whole degree vector by $\mathbf{d} = (d_1, ..., d_n)$. A (i, j)-path is a sequence of distinct vertices and edges between i and j. If a (i, j)-path exists, then i and j are said to be connected. In the end, N is connected if every pair of its vertices is connected.

The adjacency relations between nodes are represented by a $n-$square binary matrix $\mathbf{A} = [a_{ij}]$, called adjacency matrix, whose entries $a_{ij} = 1$ if $(i, j) \in E$, 0 otherwise. As we are considering an undirected network N, $\mathbf{A}$ is symmetric and its eigenvalues are real. Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ be the eigenvalues of $\mathbf{A}$, then $\mathbf{x}_i$ is the eigenvector associated to λ_i.

It is also important to specify that networks considered in this work are without loops and multiple links. More precisely, an organization can not be its own partner (i.e., there are no loops), while multiple links between nodes are considered as weights on the same link. In this vein, a network is weighted if a real number $w_{ij} > 0$ (i.e., the weight) is assigned to the link (i, j). In case of a weighted network, we denote the weighted adjacency matrix by $\mathbf{W} = [w_{ij}]$.

Finally, we say that a network is bipartite if the set V can be partitioned into two subsets V_1 and V_2 (i.e. $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$), whose cardinalities are n_1 and n_2, respectively. In a bipartite network, there are no links between nodes within V_1 and the same occurs within V_2, and every edge of N joins a node in V_1 with a node in V_2. In this case, we define the affiliation matrix $\mathbf{E}$ of order $n_1 \times n_2$, where $e_{ij} = 1$ if a link occurs between $i \in V_1$ and $j \in V_2$, 0 otherwise. Then, a bipartite network can be ultimately represented by a block square matrix of order $n_1 + n_2$, denoted by $\mathbf{B}$ (bipartite adjacency matrix) and defined as:

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{0} \end{bmatrix} \tag{2.1}$$

A bipartite network is a good representation of two-mode data, that is data representing the relations between two different sets of entities (e.g., actors and events, companies and activities, organizations and projects).

2.3.2 Network centrality measures

Centrality is one of the key issues in network analysis. Generally speaking, every element (i.e., nodes, edges, or groups of nodes) of the network can be important in respect to the entire structure, but the most investigated aspect is the assignment of a centrality score to the vertices of the network, indicating their relevance and influence in terms of connections. Among the various centrality measures existing in the literature, we focus on the family of measures based on adjacency relationships since they are the most suitable for representing the status/power of vertices. Therefore, in what follows, we introduce a subset of existing centrality measures that capture the popularity of nodes in collaborative research projects. In this vein, the most intuitive centrality measure is the degree centrality, formally represented by the degree vector $\mathbf{d} = (d_1, ..., d_n)$, introduced in Section 2.3.1. This centrality measure has an immediate interpretation in two-mode data: the degree of a node in the set V_1 (V_2, respectively) is the number of its neighbors belonging to the set V_2 (V_1, respectively).

As degree centrality, the eigenvector centrality (Bonacich, 1972) is based on adjacency relations, but with a more sophisticated interpretation. A node i is central if connected to nodes that are central themselves. In other words, the eigenvector centrality of the node i, denoted by x_i, is proportional to the sum of the centralities of its neighbors, that is:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} a_{ij} x_j, \tag{2.2}$$

where λ is a constant.

If we collect the centralities of all the nodes in the vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the expression in Eq. 2.2 can be rewritten in matrix form as $\mathbf{x} = \frac{1}{\lambda}\mathbf{A}\mathbf{x}$, so that $\mathbf{x}$ is an eigenvector of the adjacency matrix $\mathbf{A}$ with eigenvalue λ. As the centralities have to be nonnegative, by the Perron–Frobenius theorem (Horn and Johnson, 1985), the constant must be the largest eigenvalue λ_1 of $\mathbf{A}$, and the centrality vector is the corresponding principal eigenvector $\mathbf{x}_1$.

In Bonacich (1987), a generalization of the eigenvector centrality is presented by introducing the so called Beta-centrality. For a vertex $i \in V$, this measure is defined as follows:

$$c_i(\beta) = (\sum_{k=1}^{\infty} \beta^{k-1} \mathbf{A}^k \mathbf{1})_i, \tag{2.3}$$

where β is a real parameter reflecting the extent to which the status of a node is a function of the status of its neighbors. Using the vector of centrality $\mathbf{c}(\beta) = (c_1(\beta), ..., c_n(\beta)) \in \mathbb{R}^n$ we can rewrite Beta-centrality in matrix form:

$$\mathbf{c}(\beta) = \sum_{k=1}^{\infty} \beta^{k-1} \mathbf{A}^k \mathbf{1} = \frac{1}{\beta} \sum_{k=1}^{\infty} \beta^k \mathbf{A}^k \mathbf{1} \tag{2.4}$$

In Bonacich (1987), different situations have been analyzed depending on the sign of β. Nevertheless, involving the series of the $k-$power of matrix $\mathbf{A}$, β should be also an "attenuation" factor for series convergence. Thus, for our aim, and the meaning we intend to give to this centrality measure, we can assume $0 < \beta < 1$. This assumption assures the series convergence, and provides Beta-centrality with a specific meaning, i.e., a weighted sum of paths connecting any two vertices, where longer paths are weighted less. This means that the effect of neighbors' centralities on the centrality of the focal node decreases as distance increases.

Interestingly, under appropriate conditions, degree centrality, Beta-centrality, and eigenvector centrality are related, as stated by the following proposition[1]:

Proposition 1 *Let* $0 < \beta < \frac{1}{\lambda_1}$, *and* $\lambda_1 >> \lambda_i \quad (\forall i = 1, ..., n)$. *If* β *approaches* 0^+, *then* $\mathbf{c}(\beta)$ *approaches* $\mathbf{d}$*; if* β *approaches* $\frac{1}{\lambda_1}^-$, *then* $\mathbf{c}(\beta)$ *is proportional to* $\mathbf{x}_1$.

In other words, degree and eigenvector centrality can be seen as extremal cases of Beta-centrality. In particular, when β approaches 0, we recover the degree centrality, whereas when β approaches $\frac{1}{\lambda_1}$, we obtain a vector proportional to the eigenvector centrality.

2.3.3 Dual-projection approach

A variant of eigenvector centrality for two-mode data (i.e., for bipartite networks) has been proposed in Bonacich (1991). In this work, Bonacich provides not only the definition of centrality for nodes in one of the two sets, but also the idea of a mutual centrality. In general terms indeed, the centrality of an actor is proportional to the centralities of the events it is associated with. At the same time,

[1]This result can be found in (Bonacich, 2007). We reported the proof in Appendix 2.A.

the centrality of an event is proportional to the centralities of the members who participate in that event. Notably, in this work, Bonacich anticipates the dual-projection approach referring to the eigenvector centrality.

Data on participation to collaborative projects belongs to the wide category of two-mode social network data, due to its bipartite structure. Thus, relations between the sets of organizations and projects are represented by the affiliation matrix $\mathbf{E}$, whose generic element e_{ij} is equal to 1 if organization i participates in project j, and 0 otherwise. Then, the adjacency matrix $\mathbf{B}$ is defined as in Eq. (2.1).

According to Borgatti and Halgin (2011), there are two basic approaches to analyze two-mode data: the "direct" approach and the "conversion" approach. In the former case, two-mode data is analyzed without any manipulation of original data. In the latter one, two-mode data is converted into two one-mode projections and studied separately. More specifically, in the direct approach, we work with the matrix $\mathbf{B}$, while in the conversion approach, the two projection matrices $\mathbf{E}\mathbf{E}^T$ and $\mathbf{E}^T\mathbf{E}$ of order $n_1 \times n_1$ and $n_2 \times n_2$, respectively, are constructed and analyzed separately. In particular, the elements of $\mathbf{E}\mathbf{E}^T$ correspond to the number of projects shared by each pair of organizations, while the elements of $\mathbf{E}^T\mathbf{E}$ represent the number of organizations that participate in each pair of projects. Notice that, $\mathbf{E}\mathbf{E}^T$ and $\mathbf{E}^T\mathbf{E}$ are the adjacency matrices of the corresponding projected networks. The vast majority of authors, especially in the literature on the EU FPs, commonly employ the conversion approach. For this reason, we will refer to this as the "classical approach" from now on.

However, working with one-mode projections implies structural data losses, and both one-mode projections matrices $\mathbf{E}\mathbf{E}^T$ and $\mathbf{E}^T\mathbf{E}$ should be used by combining their results in the so called dual-projection approach (Everett and Borgatti, 2013). More recently, the dual-projection approach has been applied to centrality problems (Everett, 2016; Grassi et al., 2019). In this perspective, we need to define the vectors of centrality of both sets of nodes (i.e., organizations and projects) in the respective projected matrices. Then, we need to make the centrality of organizations dependent on the centrality of projects they participate in, and vice versa.

In particular, let $\mathbf{z}^{(p)} \in \mathbb{R}^{n_2}$ be the vector of a generic centrality measure associated with the nodes of the matrix $\mathbf{E}^T\mathbf{E}$, i.e., the centrality of projects obtained by the classical approach. Note that $\mathbf{z}^{(p)}$ can represent any centrality measure defined for weighted networks. The dual centrality vector of organizations $\mathbf{z}^{(o)}_{dual} \in \mathbb{R}^{n_1}$ is defined as:

$$\mathbf{z}^{(o)}_{dual} = \mathbf{E}\mathbf{z}^{(p)} \tag{2.5}$$

The same procedure can be applied to compute the dual centrality of projects. If $\mathbf{y}^{(o)} \in \mathbb{R}^{n_1}$

denotes any centrality vector of $\mathbf{EE}^T$, namely the centrality of organizations obtained by the classical approach, then the dual centrality vector $\mathbf{y}^{(p)}_{dual} \in \mathbb{R}^{n_2}$ of the projects is given by:

$$\mathbf{y}^{(p)}_{dual} = \mathbf{E}^T \mathbf{y}^{(o)} \tag{2.6}$$

Note that, by formula (2.6), the centrality of organizations obtained through the classical approach is included in the computation of the dual centrality of projects, and vice versa. In this way, it is possible to exploit the whole set of information contained in the topology of a bipartite network, thus avoiding structural data losses.

Similarly to Proposition (1), also for the dual approach, degree centrality, Beta-centrality, and eigenvector centrality are related, as stated by the following proposition[2]:

Proposition 2 *Let $\mathbf{c}(\beta)^{(p)}_{dual}$ be the dual Beta-centrality of projects, and λ_1 the largest eigenvalue of the adjacency matrix $\mathbf{EE}^T$. Assume that $0 < \beta < \frac{1}{\lambda_1}$ and $\lambda_1 >> \lambda_i \quad (\forall i = 1, ..., n)$. If β approaches 0^+, then $\mathbf{c}(\beta)^{(p)}_{dual}$ approaches the dual degree centrality of projects $\mathbf{d}^{(p)}_{dual}$; if β approaches $\frac{1}{\lambda_1}^-$, then $\mathbf{c}(\beta)^{(p)}_{dual}$ is proportional to the dual eigenvector centrality of projects $\mathbf{x}^{(p)}_{1,dual}$.*

Hence, under specific assumptions, dual degree and dual eigenvector centrality can be seen as extremal cases of dual Beta-centrality. A similar result holds for the dual Beta-centrality of organizations.

2.3.4 Classical vs dual approach: investigating different partner selection strategies

The main novelty of our approach is represented by the fact that we are able to investigate different partner selection strategies through the comparison between classical and dual centrality measures. In particular, what distiguishes the classical from the dual approach in relation to the centrality of projects is the presence of the centrality vector of organizations in the computation of the dual centrality of projects. This element enables the dual centrality of a specific project to vary based on whether an organization participating in it decides to join a new consortium with new partners rather than with existing ones. In contrast, the classical approach does not account for this distinction, and the centrality of projects remains insensitive to such different strategies.

For the sake of clarity, we provide the following example. Let us suppose that *organization r* is participating in *project i* together with *organization s*. Now, suppose that *organization r* is going to join a new consortium working on another project, that we call *project j*. At this step, we

[2] The proposition can be stated in an analogous way for the centrality of organizations. We reported the proof in Appendix 2.A.

assume the two following possibilities: in the first scenario, *organization r* decides to collaborate with *organization s* in the new project as well (i.e., it links with an existing partner); in the second scenario, *organization r* starts a new collaboration with *organization t* (i.e., it creates a link with a new partner).

In the classical approach, the centrality of projects does not capture the differences between the two different strategies. Indeed, the elements of $\mathbf{E}^T\mathbf{E}$ that change in the two scenarios coincide, as corresponding to the new adjacency relationship between *project i* and *project j*.

On the contrary, the centrality of projects reacts differently to the two strategies in the dual approach. More specifically, in the first scenario, $\mathbf{E}^T$ varies due the new affiliations of *organization s* and *organization r* to *project j*, while $\mathbf{y}^{(o)}$ remains stable since the two organizations were already connected in *project i*. In the second scenario instead, besides variations in $\mathbf{E}^T$ because of the new affiliations of *organization s* and *organization r* to *project j*, also the centrality vector of organizations $\mathbf{y}^{(o)}$ changes as a consequence of the new connection between *organization r* and *organization t*. In particular, this difference is significantly relevant in magnitude for unweighted centrality measures. This is why we consider degree centrality, eigenvector centrality, and Beta centrality in our analysis, thereby disregarding edge weights.

Therefore, by comparing the effects of classical and dual centrality measures on the economic contribution received by the project, we are able to analyze whether it is better to collaborate with new or existing partners to be more likely to obtain a higher amount of funds. A graphic representation of the aforementioned dynamics, along with their effect on the centrality of projects in both approaches, is reported in Figure 2.1.

2.4 Sample and data

The analysis of projects funded by the EU FPs represents an intriguing case study in the field of collaborative R&D at the European level. In particular, understanding the way in which the configuration of participants' connections in a collaborative network affects the amount of funds awarded to a research project is extremely relevant both from a policy perspective and for the collaborative strategies of participants.

We take into account projects funded by the eighth FP, H2020. H2020 was the EU research and innovation funding programme from 2014 to 2020 with a budget of nearly 80 billion €. It was implemented, as all the EU FPs, via multi-annual work programmes aimed at addressing many of the European Commission's policy objectives. Specifically, H2020 focuses on three main priorities: excellent science, industrial leadership, and societal challenges (European Commission, 2014). Data on projects funded by H2020 are publicly available from the Community Research and Develop-

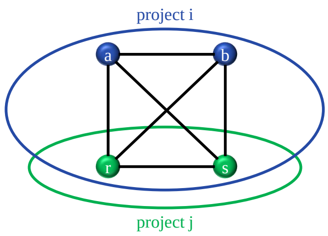
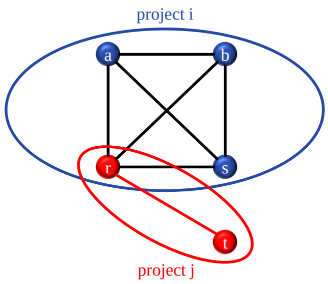
Classical approach $y_i^{(p)}$ depends on $E^T E$	**Partner selection strategies**	**Dual approach** $y_{i,dual}^{(p)} = E^T y^{(o)}$
Scenario 1: Adjacency relations of $E^T E$ change due to the link between *project i* and *project j*.		**Scenario 1:** E^T varies due to the affiliation of *organization r* to *project j*; $y^{(o)}$ does not vary.
Scenario 2: Adjacency relations of $E^T E$ change as above due to the link between *project i* and *project j*.		**Scenario 2:** E^T varies due to the affiliation of *organization r* to *project j*; $y^{(o)}$ varies due to the link between *organization r* and *organization t*.

Figure 2.1: Classical vs dual approach: representation of two different partner selection strategies. Scenario 1: linking with an existing partner. Scenario 2: linking with a new partner. Nodes represent organizations, and links stand for collaboration in the same project; then, networks represent projects.

ment Information Service (CORDIS) website[3]. However, raw data from CORDIS website comprises non-standardized organization names, as well as possible sources of errors (Ancona et al., 2023a). Roediger-Schluga and Barber (2008) tackle this issue, disambiguating participant organizations through a proper labor-intensive procedure relying on a consistent manual inspection of participants. In this way, the authors obtain a novel data source of higher quality than the original data, known by the name of "EUPRO database". Actually, most of the studies dealing with European projects benefit from the EUPRO database, performing their analyses on this data source (Paier and Scherngell, 2011; Scherngell and Barber, 2011; Hoekman et al., 2013; Scherngell and Lata, 2013; Lepori et al., 2015; Heringa et al., 2016; Cavallaro and Lepori, 2021).

Thus, we requested and obtained access to the EUPRO database, restricted to the eighth EU FP. This data includes information on 173,084 participations in 35,378 projects. In particular, the number of distinct organizations is equal to 40,470. However, in order to determine the subject of the analysis, we must identify those members belonging to the giant component of the bipartite network, i.e., the maximal connected component of the network. This step is indeed necessary to compute the eigenvector centrality and the Beta-centrality of all projects and organizations. Then, we generate the two projected networks in R (R Core Team, 2014) to determine the vectors of centrality scores of both organizations and projects to be used in the classical and dual approach. The projected network of organizations (called "participant network" from now on) is an undirected

[3]https://cordis.europa.eu/

weighted network whose nodes represent participants in projects, edges stand for collaborations between the nodes they connect, and weights on edges are equal to the number of projects in which the respective organizations collaborate. The projected network of projects (called "project network" from now on) is an undirected weighted network whose nodes represent projects, there is an edge connecting two nodes if they share at least one participant, and weights on edges are equal to the number of organizations participating in both the related projects. The participant network is constituted by $40,470$ nodes and $1,074,618$ edges, while the project network is made by $30,382$ nodes and $9,643,357$ edges. The two networks have also some attributes associated with their nodes, namely organization name, country, and activity type for the participant network, and funding scheme, coordinator country, and coordinator activity type for the project network. We recall that the subject of our analysis are projects, but the attributes of both networks will be employed in our econometric analysis.

Since EUPRO data does not include details about the economic contribution received by the projects from the European Commission, we had to retrieve this information from the CORDIS dataset of H2020. Therefore, the final sample of projects used for our analysis is identified in correspondence with the intersection between the EUPRO database and the CORDIS dataset in relation to the eighth EU FP. Specifically, the number of projects reporting all necessary data for the analysis is equal to $29,470$. It is important to specify that centrality measures are computed for all the projects belonging to the project network introduced above in order not to alter the effective connection structure of projects, although the values of centrality measures entering the model are only those related to the subset of projects here defined.

2.5 Econometric model

To investigate the effects of partner selection strategies on receiving a higher amount of funds in H2020-funded projects, we run a multiple linear regression analysis. While we acknowledge that testing a linear model between centrality measures and economic contributions can be a strong assumption, we want to specify that the main objective of this research is to compare the effects of different metrics computed through classical and dual approach, respectively, on the economic contribution received by the project. This comparison aims to determine which partner selection strategy is likely to provide the project with a higher amount of funding. In particular, we analyze five different specifications, each implementing a specific centrality measure as a proxy of the connection structure of projects, i.e., degree centrality through classical approach, and eigenvector centrality and Beta-centrality through both classical and dual approaches. The model comprises both numerical and categorical variables. Specifically, all numerical variables are standardized in order to interpret the results. The approach is similar to that proposed in (Ebadi and Schiffauerova, 2015), where the authors analyze the impact of collaboration and networking on the amount of funding received by Canadian researchers. However, there are three main differences that distinguish their study from our work. First, their analysis is performed at the individual level, while we investigate the effects of collaboration at the organizational level. Second, we focus on projects funded by H2020, which, given its relevance and dimension in terms of funded projects, represents one of the most interesting case studies in the field of collaborative R&D. Third, we also consider centrality measures in their dual form, whereas Ebadi and Schiffauerova (2015) implement degree centrality, betweenness centrality, eigenvector centrality, and clustering coefficient only through the classical approach.

Dependent variable

Our dependent variable is the economic contribution received by the projects from the European Commission (*ecContr* in the model). This data is retrieved from the CORDIS dataset.

Explanatory variables

The independent variables of interest in this study are the three centrality measures introduced in Section 2.3.2. Each of them is able to capture a different characteristic of the connection structure of projects, and, consequently, a different behavior of the nodes of the project network. Degree centrality ($d^{(p)}$ in the model) represents the number of connections a node establishes in the network. Specifically, a project is linked to another project if they share at least one common participant. Thus, the more an organization is involved in European projects, the higher the degree of the projects it takes part in. Eigenvector centrality through classical approach ($x_1^{(p)}$ in the model) considers not only how many neighbors (i.e., links) a project has, but also the quality (in terms of centrality) of its neighbors' connections. Thus, it goes more in depth compared to degree centrality to proxy the connection structure of nodes. Similarly, Beta-centrality through classical approach ($c_i(\beta)^{(p)}$ in the model) takes into account the connection structure of the project's neighborhood, but it introduces a scaling factor weighting more the connections of closer neighbors. Finally, eigenvector centrality through dual approach ($x_{1,dual}^{(p)}$ in the model) and Beta-centrality through dual approach ($c_i(\beta)_{dual}^{(p)}$ in the model) are the centrality measures that enable us to compare different partner selection strategies, distinguishing between the connection with new and existing partners. In fact, as described in Section 2.3.4, the dual-projection approach allows accounting for the centrality of organizations in the computation of the centrality of projects. This means that, additionally to the classical approach, eigenvector and Beta-centrality through dual approach also considers the relevance (in terms of the connection structure of organizations in the participant network) of participants in linking two projects. It is important to specify that, since Beta-centrality is a parametric function, we need to fix the value of β to analyze its effect on the economic contribution. In line with (Everett, 2016), we consider values of Beta-centrality in correspondence with $\beta = \frac{1}{2\lambda_1}$, i.e., the midpoint between 0 and β_{max}.

Control variables

Based on the literature reported in Section 2.2, we control for the effect of other independent variables on the economic contribution. We assume that those factors that are considered important to determine project outcomes, or increase the probability of being funded, may reasonably have an impact on the amount of funds as well (taking into account variables we can control for). In particular, we include one numerical variable and three categorical variables in the model that may affect the amount of funds received by a project. The numerical variable represents the level of diversity of a project (Gattringer et al., 2017; Wu et al., 2019). We measure the project diversity through the Blau index (Blau, 1977) by country (*div_country* in the model) and activity type (*div_actType* in the model), following previous works in the literature using the Blau index as a proxy of top management team diversity (Nielsen and Nielsen, 2013), portfolio diversity (Jiang et al., 2010), and partner diversity (Alonso and Andrews, 2019). The Blau index is defined as $B = 1 - \sum p_i^2$, where, in our case, B represents the degree of project diversity, and p_i corresponds to the proportion of organizations participating in the project that belong to category i.

For what concerns the categorical variables, we consider two variables related to the characteristics of the project coordinator, namely the activity type (*actType* in the model) (Enger, 2018; Balland et al., 2019) and the country (*macroCountry* in the model) (Balland et al., 2019; Cavallaro and Lepori, 2021). The activity type indicates if the coordinator is a private company (PRC), a higher education institution (HES), a public body (PUB), a research organization (REC), or other (OTH). Regarding the latter variable instead, we followed the approach of Balland et al. (2019), hence aggregating countries in macro-categories, i.e., "EU-15", "EU-13", and "Extra-EU". For completeness, we also run regressions on single countries, but they did not provide significant results[4]. Then, we control for the funding scheme ($fundScheme$ in the model) under which the project has been financed, as we believe that it can plausibly affect the economic contribution awarded to the projects. For each categorical variable, $n-1$ dummy variables are generated, where n is the number of classes of the related variable. In the dataset there are 5 activity types, 3 macro-countries, and 64 funding schemes, thus 69 dummy variables are introduced. Including this considerable amount

[4]Results are available from the authors upon request.

of variables did not lead to overfitting issues due to the large number of observations we have in the model, i.e., the degree of freedoms are still remarkable.

The final specifications are defined as follows:

Specification I

$$\begin{aligned} ecContr_i = \alpha_0^{(I)} + \alpha_1^{(I)} x_{1,dual,i}^{(p)} + \alpha_2^{(I)} div\_country_i + \alpha_3^{(I)} div\_actType_i + \alpha_4^{(I)} macroCountry1_i + \\ + \alpha_5^{(I)} macroCountry2_i + \alpha_6^{(I)} actType1_i + ... + \alpha_9^{(I)} actType4_i + \alpha_{10}^{(I)} fundScheme1_i + \\ + ... + \alpha_{72}^{(I)} fundScheme63_i \end{aligned}$$

Specification II

$$\begin{aligned} ecContr_i = \alpha_0^{(II)} + \alpha_1^{(II)} x_{1,i}^{(p)} + \alpha_2^{(II)} div\_country_i + \alpha_3^{(II)} div\_actType_i + \alpha_4^{(II)} macroCountry1_i + \\ + \alpha_5^{(II)} macroCountry2_i + \alpha_6^{(II)} actType1_i + ... + \alpha_9^{(II)} actType4_i + \alpha_{10}^{(II)} fundScheme1_i + \\ + ... + \alpha_{72}^{(II)} fundScheme63_i \end{aligned}$$

Specification III

$$\begin{aligned} ecContr_i = \alpha_0^{(III)} + \alpha_1^{(III)} c_i(\beta)_{dual}^{(p)} + \alpha_2^{(III)} div\_country_i + \alpha_3^{(III)} div\_actType_i + \alpha_4^{(III)} macroCountry1_i + \\ + \alpha_5^{(III)} macroCountry2_i + \alpha_6^{(III)} actType1_i + ... + \alpha_9^{(III)} actType4_i + \alpha_{10}^{(III)} fundScheme1_i + \\ + ... + \alpha_{72}^{(III)} fundScheme63_i \end{aligned}$$

Specification IV

$$\begin{aligned} ecContr_i = \alpha_0^{(IV)} + \alpha_1^{(IV)} c_i(\beta)^{(p)} + \alpha_2^{(IV)} div\_country_i + \alpha_3^{(IV)} div\_actType_i + \alpha_4^{(IV)} macroCountry1_i + \\ + \alpha_5^{(IV)} macroCountry2_i + \alpha_6^{(IV)} actType1_i + ... + \alpha_9^{(IV)} actType4_i + \alpha_{10}^{(IV)} fundScheme1_i + \\ + ... + \alpha_{72}^{(IV)} fundScheme63_i \end{aligned}$$

Specification V

$$\begin{aligned} ecContr_i = \alpha_0^{(V)} + \alpha_1^{(V)} d_i^{(p)} + \alpha_2^{(V)} div\_country_i + \alpha_3^{(V)} div\_actType_i + \alpha_4^{(V)} macroCountry1_i + \\ + \alpha_5^{(V)} macroCountry2_i + \alpha_6^{(V)} actType1_i + ... + \alpha_9^{(V)} actType4_i + \alpha_{10}^{(V)} fundScheme1_i + \\ + ... + \alpha_{72}^{(V)} fundScheme63_i \end{aligned}$$

Where the generic $\alpha^{(S)}$'s are the coefficients of the variables to be estimated in Specification S.

In light of the aforementioned specifications, we express our hypotheses in terms of the variables defined in this section. Regarding H1, we consider a *p-value* < 0.001 as strong evidence against the null hypothesis, consistent with classical econometric tests.

Hypothesis 1 *p-value*($\alpha_1^{(I)}$), *p-value*($\alpha_1^{(II)}$), *p-value*($\alpha_1^{(III)}$), *p-value*($\alpha_1^{(IV)}$), *p-value*($\alpha_1^{(V)}$) < 0.001

Hypothesis 2 $\alpha_1^{(I)} < \alpha_1^{(II)}$, *and* $\alpha_1^{(III)} < \alpha_1^{(IV)}$

Hypothesis 3 $\alpha_2^{(I)}, \alpha_2^{(II)}, \alpha_2^{(III)}, \alpha_2^{(IV)}, \alpha_2^{(V)} > 0$, *and* $\alpha_3^{(I)}, \alpha_3^{(II)}, \alpha_3^{(III)}, \alpha_3^{(IV)}, \alpha_3^{(V)} > 0$

Hypothesis 4 $\alpha_1^{(I)}, \alpha_1^{(II)}, \alpha_1^{(III)}, \alpha_1^{(IV)}, \alpha_1^{(V)} > 0$

2.6 Results

In Table 2.1, we report the descriptive statistics of the dependent variable and the explanatory variables included in the different specifications. Looking at these, the overall distribution of economic contribution seems to be skewed, thereby potentially undermining one of the assumptions of linear regressions. However, this is due to the presence of outliers in the distribution. In fact, if we look at the kernel density of the economic contribution among the set of projects belonging to the giant component (in Figure 2.2 below), the normal distribution assumption of the dependent variable appears more reasonable. Moreover, as specified in Section 2.5, our primary goal is not to find a perfectly fitting econometric model, but rather to compare classic and dual centrality measures in a more refined and comprehensive manner than through correlation coefficients.

Table 2.1: Descriptive statistics.

	Economic contribution	Dual eigenvector	Classic eigenvector	Dual Beta ($\beta = \frac{1}{2\lambda_1}$)	Classic Beta ($\beta = \frac{1}{2\lambda_1}$)	Degree
Mean	€$2,176,239.00$	0.49	0.02	25.46	0.59	636.46
SD	€$5,877,372.00$	0.84	0.05	40.93	0.81	866.62
Min	€$3,150.00$	8.05×10^{-8}	1.60×10^{-9}	0.01	5.98×10^{-4}	1
Max	€$678,800,000.00$	25.30	1	1130.88	11.84	$15,039$

In Figure 2.3, we represent the Kendall τ correlation matrix comparing the rankings provided by the different metrics. As it is well known in SNA, all centrality measures are mutually highly correlated (Valente et al., 2008), and this aspect is supported by the obtained results. However, the Kendall τ correlation matrix highlights a first interesting outcome of our analysis. The rankings provided by the centrality measures through the dual approach show a higher correlation with the ranking based on the economic contribution compared to the rankings provided by the same metrics through the classical approach. More specifically, dual Beta-centrality is the most highly correlated with the economic contribution, while, interestingly, eigenvector centrality overperforms Beta-centrality in their classical version. These results highlight two first interesting aspects. First, the dual-projection method allows to better identify the projects according to the size of their

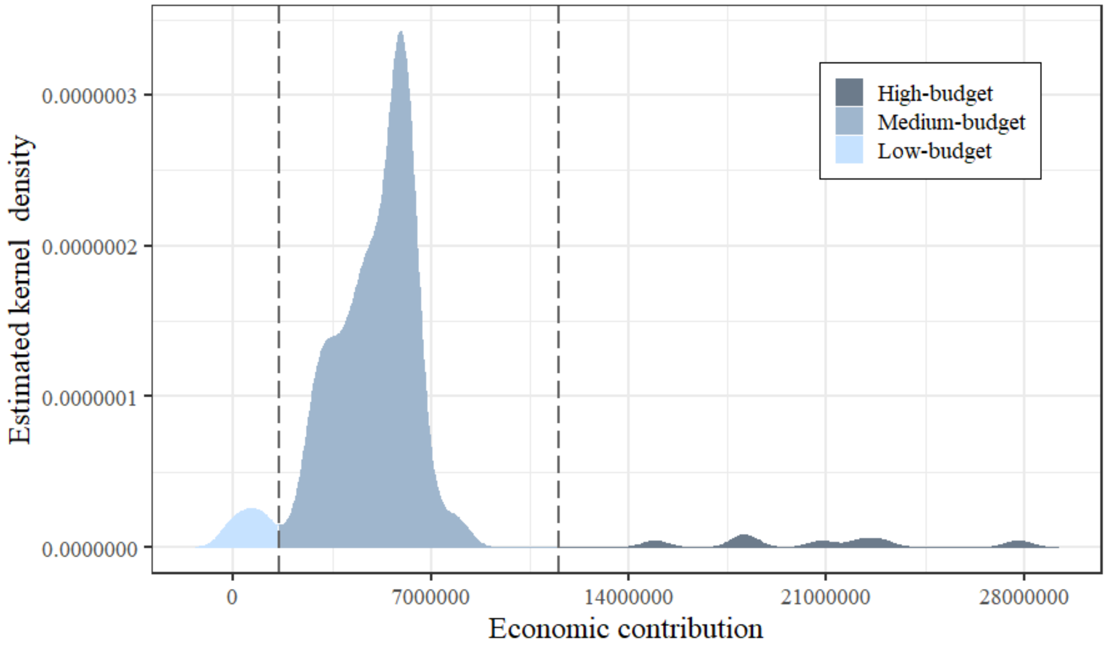


Figure 2.2: Kernel density distribution of the economic contribution received by the projects belonging to the giant component. The thresholds for high-, medium-, and low-budget projects are identified in correspondence with local minima, and are displayed for illustrative purposes.

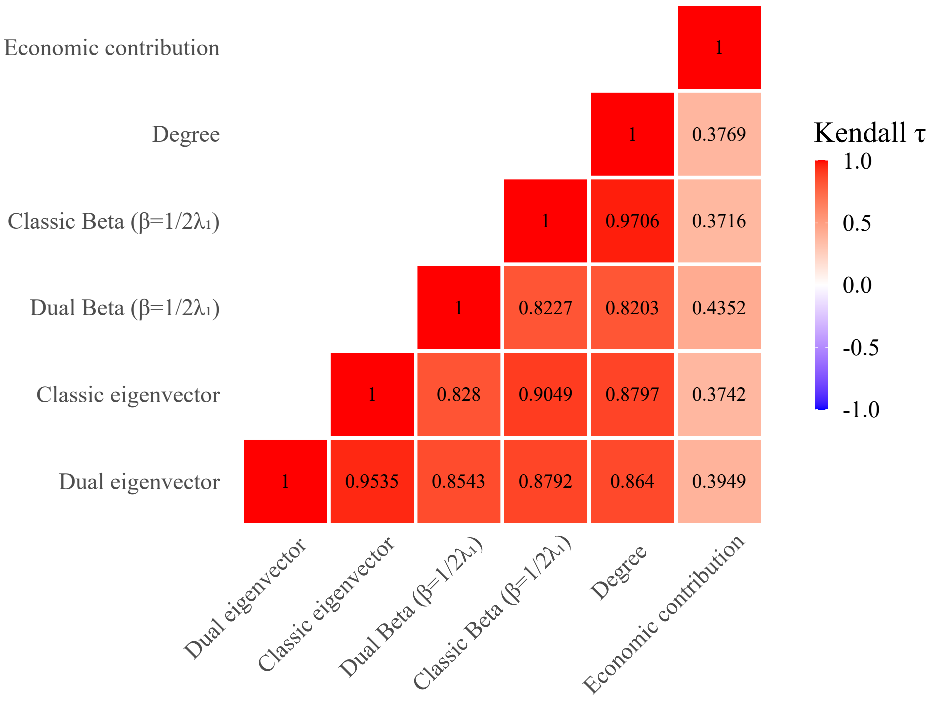


Figure 2.3: Kendall τ correlation matrix.

financial support. Second, dual Beta-centrality seems to be the best proxy (among the ones we tested) to estimate the amount of received funds by H2020 funded projects.

Table 2.2 presents the results obtained from the multiple linear regression described in Section 2.5. We report robust standard errors using the Huber-White sandwich estimator (White, 1980) in order to correct for heteroskedasticity and avoid any potential effects of non-independent observations. Our model estimations are not affected by multicollinearity issues. The individual variance inflation factors (VIFs) are indeed below the recommended cutoff level of 10 for all variables in the different specifications (the maximum value is equal to 2.58) (Neter et al., 1996).

Table 2.2: Results from the econometric analysis. Robust standard errors are in parentheses.

	(I) Economic contribution	(II) Economic contribution	(III) Economic contribution	(IV) Economic contribution	(V) Economic contribution
Intercept	-0.234^{*}	-0.276^{***}	-0.231^{*}	-0.299^{***}	-0.276^{***}
	(0.094)	(0.079)	(0.108)	(0.074)	(0.080)
Dual eigenvector	0.390^{***}				
	(0.073)				
Classic eigenvector		0.280^{***}			
		(0.056)			
Dual Beta ($\beta = \frac{1}{2\lambda_1}$)			0.468^{***}		
			(0.088)		
Classic Beta ($\beta = \frac{1}{2\lambda_1}$)				0.241^{***}	
				(0.039)	
Degree					0.281^{***}
					(0.046)
Diversity actType	-0.047^{**}	-0.043^{**}	-0.061^{***}	-0.042^{**}	-0.044^{**}
	(0.015)	(0.015)	(0.017)	(0.014)	(0.014)
Diversity country	0.070^{*}	0.125^{***}	0.012	0.119^{***}	0.098^{***}
	(0.030)	(0.024)	(0.038)	(0.022)	(0.024)
ActType OTH	0.106^{***}	0.058^{**}	0.139^{***}	0.079^{***}	0.099^{***}
	(0.026)	(0.022)	(0.031)	(0.023)	(0.024)
ActType PRC	0.119^{***}	0.064^{*}	0.147^{***}	0.080^{**}	0.105^{***}
	(0.030)	(0.026)	(0.034)	(0.026)	(0.028)
ActType PUB	0.167^{**}	0.130^{*}	0.183^{***}	0.146^{**}	0.160^{**}
	(0.052)	(0.054)	(0.050)	(0.054)	(0.053)
ActType REC	-0.063^{***}	-0.089^{***}	-0.035^{**}	-0.057^{***}	-0.048^{***}
	(0.016)	(0.022)	(0.011)	(0.014)	(0.013)
Macrocountry EU-15	-0.058^{***}	-0.041^{**}	-0.070^{***}	-0.046^{***}	-0.056^{***}
	(0.016)	(0.014)	(0.018)	(0.013)	(0.015)
Macrocountry Extra-EU	-0.033	-0.010	-0.027	-0.007	-0.016
	(0.019)	(0.017)	(0.019)	(0.017)	(0.017)
FundScheme BBI-IA-DEMO	1.088^{***}	1.086^{***}	1.080^{***}	1.102^{***}	1.105^{***}
	(0.108)	(0.099)	(0.124)	(0.100)	(0.103)
FundScheme BBI-IA-FLAG	3.400^{***}	3.401^{***}	3.427^{***}	3.407^{***}	3.402^{***}
	(0.212)	(0.205)	(0.223)	(0.205)	(0.208)
FundScheme BBI-RIA	0.376^{***}	0.401^{***}	0.320^{**}	0.389^{***}	0.384^{***}
	(0.082)	(0.068)	(0.102)	(0.069)	(0.074)
FundScheme COFUND	0.762^{***}	0.846^{***}	0.716^{***}	0.910^{***}	0.921^{***}
	(0.083)	(0.064)	(0.103)	(0.062)	(0.066)
FundScheme COFUND-EJP	18.299	20.019	16.890	20.820	20.407
	(13.013)	(13.659)	(12.516)	(14.070)	(13.935)
FundScheme COFUND-PCP	0.532^{***}	0.497^{***}	0.611^{***}	0.518^{***}	0.537^{***}
	(0.090)	(0.093)	(0.101)	(0.089)	(0.088)
FundScheme CS2-CSA	0.118	0.041	0.201	0.067	0.087
	(0.102)	(0.114)	(0.114)	(0.105)	(0.106)

Continued on next page

Table 2.2 – *Continued from previous page*

	(I) Economic contribution	(II) Economic contribution	(III) Economic contribution	(IV) Economic contribution	(V) Economic contribution
FundScheme CS2-IA	0.865***	0.888***	0.870***	0.905***	0.896***
	(0.262)	(0.263)	(0.262)	(0.264)	(0.264)
FundScheme CS2-RIA	0.209**	0.218***	0.233*	0.232***	0.228**
	(0.078)	(0.065)	(0.094)	(0.066)	(0.070)
FundScheme CSA	0.105	0.139*	0.104	0.153*	0.142*
	(0.075)	(0.059)	(0.092)	(0.060)	(0.065)
FundScheme CSA-LS	0.033	0.004	0.097	0.022	0.033
	(0.079)	(0.069)	(0.100)	(0.070)	(0.073)
FundScheme CSA-LSP	0.053	0.073	0.063	0.093	0.085
	(0.084)	(0.071)	(0.100)	(0.069)	(0.074)
FundScheme ECSEL-CSA	−0.269	−0.186	−0.306	−0.142	−0.151
	(0.243)	(0.185)	(0.269)	(0.162)	(0.179)
FundScheme ECSEL-IA	1.679***	2.045***	1.219*	2.158***	2.064***
	(0.419)	(0.370)	(0.504)	(0.356)	(0.365)
FundScheme ECSEL-RIA	0.320	0.631***	−0.050	0.720***	0.644***
	(0.194)	(0.153)	(0.266)	(0.134)	(0.142)
FundScheme ERA-NET-Cofund	0.591***	0.756***	0.540***	0.868***	0.866***
	(0.107)	(0.083)	(0.125)	(0.081)	(0.084)
FundScheme ERC-ADG	0.456***	0.479***	0.471***	0.495***	0.479***
	(0.093)	(0.081)	(0.106)	(0.078)	(0.083)
FundScheme ERC-COG	0.382***	0.403***	0.397***	0.420***	0.406***
	(0.093)	(0.081)	(0.106)	(0.078)	(0.083)
FundScheme ERC-LVG	0.090	0.145	0.080	0.163*	0.145
	(0.092)	(0.079)	(0.108)	(0.076)	(0.081)
FundScheme ERC-POC	0.108	0.141	0.113	0.156*	0.141
	(0.092)	(0.080)	(0.106)	(0.078)	(0.082)
FundScheme ERC-POC-LS	0.110	0.138	0.111	0.157*	0.144
	(0.093)	(0.081)	(0.107)	(0.079)	(0.083)
FundScheme ERC-STG	0.311***	0.334***	0.323**	0.350***	0.336***
	(0.093)	(0.081)	(0.106)	(0.078)	(0.083)
FundScheme ERC-SyG	1.496***	1.493***	1.593***	1.509***	1.495***
	(0.093)	(0.085)	(0.102)	(0.081)	(0.086)
FundScheme EuroHPC-IA	−0.072	0.095	−0.033	0.138	0.116
	(0.098)	(0.076)	(0.107)	(0.075)	(0.075)
FundScheme EuroHPC-RIA	0.592	0.580	0.658	0.587	0.593
	(0.460)	(0.440)	(0.490)	(0.463)	(0.480)
FundScheme FCH2-CSA	−0.213	−0.112	−0.183	−0.077	−0.086
	(0.141)	(0.102)	(0.165)	(0.100)	(0.112)
FundScheme FCH2-IA	1.612***	1.628***	1.615***	1.635***	1.638***
	(0.454)	(0.449)	(0.450)	(0.453)	(0.454)
FundScheme FCH2-RIA	0.154	0.195*	0.187	0.224**	0.218**
	(0.093)	(0.077)	(0.106)	(0.073)	(0.078)
FundScheme H2020-EEN-SGA	0.151	0.188**	0.157	0.210**	0.196*
	(0.082)	(0.070)	(0.098)	(0.072)	(0.076)
FundScheme IA	0.823***	0.877***	0.729***	0.880***	0.863***
	(0.079)	(0.062)	(0.102)	(0.063)	(0.068)
FundScheme IA-LS	0.719***	0.890***	0.347*	0.856***	0.804***
	(0.099)	(0.068)	(0.165)	(0.070)	(0.079)

Continued on next page

Table 2.2 – *Continued from previous page*

	(I) Economic contribution	(II) Economic contribution	(III) Economic contribution	(IV) Economic contribution	(V) Economic contribution
FundScheme IMI2-CSA	0.058	0.160	−0.044	0.205*	0.192**
	(0.079)	(0.085)	(0.129)	(0.081)	(0.071)
FundScheme IMI2-RIA	1.342***	1.536***	1.253**	1.538***	1.468***
	(0.383)	(0.388)	(0.384)	(0.388)	(0.387)
FundScheme MSCA-COFUND	0.583***	0.628***	0.576***	0.630***	0.609***
	(0.134)	(0.124)	(0.143)	(0.128)	(0.132)
FundScheme MSCA-COFUND-DP	0.329**	0.384***	0.310**	0.406***	0.385***
	(0.104)	(0.088)	(0.119)	(0.084)	(0.090)
FundScheme MSCA-COFUND-FP	0.554***	0.612***	0.547***	0.626***	0.603***
	(0.100)	(0.084)	(0.114)	(0.083)	(0.089)
FundScheme MSCA-IF	0.098	0.119	0.110	0.123	0.106
	(0.089)	(0.076)	(0.103)	(0.075)	(0.080)
FundScheme MSCA-IF-EF-CAR	0.137	0.168*	0.131	0.186*	0.171*
	(0.092)	(0.079)	(0.107)	(0.078)	(0.082)
FundScheme MSCA-IF-EF-RI	0.110	0.133	0.110	0.153	0.140
	(0.094)	(0.082)	(0.108)	(0.079)	(0.084)
FundScheme MSCA-IF-EF-SE	0.063	0.120	0.044	0.142*	0.118
	(0.087)	(0.070)	(0.104)	(0.068)	(0.074)
FundScheme MSCA-IF-EF-ST	0.117	0.147	0.115	0.161*	0.145
	(0.093)	(0.081)	(0.107)	(0.078)	(0.083)
FundScheme MSCA-IF-GF	0.058	0.012	0.132	0.034	0.046
	(0.074)	(0.061)	(0.092)	(0.063)	(0.067)
FundScheme MSCA-ITN	0.090	0.178*	0.119	0.162*	0.107
	(0.103)	(0.083)	(0.113)	(0.079)	(0.089)
FundScheme MSCA-ITN-EID	0.173*	0.166**	0.230*	0.176**	0.179**
	(0.074)	(0.059)	(0.092)	(0.061)	(0.066)
FundScheme MSCA-ITN-EJD	0.225*	0.283***	0.233*	0.289***	0.268***
	(0.093)	(0.073)	(0.108)	(0.071)	(0.078)
FundScheme MSCA-ITN-ETN	0.061	0.161	0.066	0.149	0.092
	(0.106)	(0.085)	(0.118)	(0.079)	(0.090)
FundScheme MSCA-RISE	−0.040	−0.056	0.012	−0.042	−0.040
	(0.073)	(0.059)	(0.090)	(0.060)	(0.065)
FundScheme PCP	0.991***	0.963***	1.049***	0.990***	1.006***
	(0.122)	(0.112)	(0.129)	(0.109)	(0.112)
FundScheme PPI	0.493*	0.420*	0.602**	0.454*	0.485*
	(0.200)	(0.205)	(0.202)	(0.200)	(0.196)
FundScheme RIA	0.431***	0.528***	0.400***	0.553***	0.515***
	(0.096)	(0.073)	(0.114)	(0.067)	(0.075)
FundScheme RIA-LS	0.612**	0.626***	0.690**	0.661***	0.638***
	(0.208)	(0.182)	(0.213)	(0.172)	(0.184)
FundScheme SESAR-CSA	0.360*	0.457***	0.337*	0.505***	0.512***
	(0.145)	(0.133)	(0.167)	(0.136)	(0.140)
FundScheme SESAR-IA	0.470***	0.499***	0.478***	0.517***	0.516***
	(0.133)	(0.129)	(0.141)	(0.132)	(0.135)
FundScheme SESAR-RIA	0.502***	0.545***	0.494***	0.562***	0.560***
	(0.110)	(0.107)	(0.119)	(0.110)	(0.112)
FundScheme Shift2Rail-CSA	0.119	0.133	0.088	0.122	0.113
	(0.095)	(0.075)	(0.123)	(0.076)	(0.082)

Continued on next page

Table 2.2 – *Continued from previous page*

	(I) Economic contribution	(II) Economic contribution	(III) Economic contribution	(IV) Economic contribution	(V) Economic contribution
FundScheme Shift2Rail-IA	0.358***	0.355**	0.288**	0.335***	0.329***
	(0.103)	(0.126)	(0.097)	(0.097)	(0.084)
FundScheme Shift2Rail-IA-LS	1.678**	1.733*	1.587**	1.759*	1.756*
	(0.645)	(0.713)	(0.561)	(0.714)	(0.701)
FundScheme Shift2Rail-RIA	0.221*	0.238**	0.212*	0.235**	0.230**
	(0.091)	(0.078)	(0.106)	(0.081)	(0.085)
FundScheme Shift2Rail-RIA-LS	1.360**	1.337**	1.386**	1.379**	1.392**
	(0.492)	(0.486)	(0.490)	(0.488)	(0.491)
FundScheme SME	0.327***	0.372***	0.324**	0.399***	0.380***
	(0.090)	(0.076)	(0.105)	(0.074)	(0.079)
FundScheme SME-1	0.046	0.101	0.033	0.127	0.104
	(0.088)	(0.072)	(0.104)	(0.069)	(0.075)
FundScheme SME-2	0.332***	0.385***	0.322**	0.412***	0.389***
	(0.088)	(0.072)	(0.104)	(0.069)	(0.075)
FundScheme SME-2b	0.363***	0.416***	0.352***	0.442***	0.420***
	(0.089)	(0.073)	(0.105)	(0.071)	(0.076)
Residual Standard Error	0.7910	0.8186	0.7760	0.8302	0.8233
Degrees of freedom	29,397	29,397	29,397	29,397	29,397
Adjusted R-squared	0.3744	0.3316	0.3978	0.3107	0.3221
F-statistic	245.9	202.5	271.4	185.5	195.5

Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

As evident from the results, all the centrality measures tested within the model exhibit a positive and significant relationship with the economic contribution received by the project. This finding confirms that the network structure of projects plays a crucial role in determining the amount of obtained funds (H1, H4). More specifically, the dual Beta-centrality has the highest coefficient among all centrality measures, and the corresponding specification is the one that best fits the model (as shown by the values of adjusted R-squared and F-statistic). Thus, by capturing the significance of participant organizations in the connection structure of projects, more accurate results can be obtained when estimating the effects on the economic contribution. This implies that it is not only important for a project to have a high number of links with other projects through shared organizations, but it is also crucial to consider the specific organizations that make these connections happen. As described in Section 2.3.4, the dual projection method is sensitive to the selection of new partners, while centrality measures determined by classical approach do not entail the difference between new and existing partners. Therefore, our results suggest that selecting new over existing partners when entering a new project increases the likelihood of obtaining more funds in European research projects. This is a relevant insight from our analysis, providing new evidence on this aspect in the collaboration literature, and violating H2. Nevertheless, this result can be seen as an indication of the relevance of enhancing the organization's social capital (Seo, 2020) by linking with partners that it has not previously collaborated with. For what concerns H3 instead, we find that the diversity of participants is positively related with the economic contribution based on the country of the organizations, while the effect is the opposite in terms of activity type heterogeneity. However, both values are relatively small in magnitude, and country diversity, in particular, is not significant within the specification including dual Beta-centrality as independent variable. Another interesting result concerns the effect of the activity type of the coordinator on the economic contribution. In particular, values for private companies, public bodies, and research organizations are significant in all five specifications with different signs. More specifically, having a private company or a public organization as a project coordinator increases the likelihood of getting more funds. On the contrary, consortia with a research institution as a coordinator decrease their probability of receiving a high amount of funds (although the effect is not big in magnitude). Results are quite surprising regarding projects with a coordinator from a EU-15 country. Indeed, such projects are less likely to obtain high funding. Finally, the funding scheme with the highest significant impact on economic contribution for all three specifications is BBI-IA-FLAG, which corresponds to “Bio-based industries Innovation Action Flagship” within the H2020 pillar “Industrial Leadership - Leadership in enabling and industrial technologies – Biotechnology”.

2.7 Discussion and conclusions

In this work, we contribute to the existing body of literature that explores the impact of partner selection strategies on the amount of funds in the field of collaborative research projects. In particular, we fill a relevant gap in the collaboration literature, which lacks evidence on the analysis of strategic choices behind partner selection mechanisms. To this aim, we test four hypotheses which are mostly derived and adapted from the analysis of partner selection strategies in the alliance literature. We also extend the current stream of research on the analysis of projects funded by the EU FPs through the investigation of a specific case study represented by the projects funded by the eighth EU FP, i.e., H2020.

The main peculiarity of this study is represented by the application of the dual-projection approach (Everett and Borgatti, 2013) to collaborative research networks. Thanks to the implementation of this novel methodology, we can express the centrality of projects, i.e., the level of importance based on the number of their connections, as a function of the centrality of organizations. In this way, it is possible to exploit the whole set of information included in the topology of a bipartite network, avoiding structural data losses occurring when working with one-mode projections, as it commonly happens in collaborative network studies. In doing so, we can investigate whether it is better (in terms of obtained funds) to join consortia by partnering with previously unrelated organizations or collaborating with already existing partners. To this aim, we compare the effects of different centrality measures through classical and dual approach on the amount of funds received by the projects. In particular, we find that the connection structure (in terms of centrality measures) of projects funded by H2020 has a positive and significant impact on the economic contribution they receive. Thus, the more a project assumes a central position within the collaborative network, sharing participants with a large number of other projects, the higher the probability of securing a significant amount of funds. This result is in line with previous studies both in the alliance (Weng et al., 2014; Kang and Zaheer, 2018; Kumar and Zaheer, 2019), and in the EU FPs (Enger, 2018; Balland et al., 2019; Cinelli et al., 2022) literature.

However, the main contribution of this study lies in the comparison between classical and dual centrality measures to investigate different partner selection strategies. Our results indeed, shed light on a controversial aspect of the collaboration literature regarding the predominant role of social capital (Seo, 2020) compared to increasing the level of trust (Bruneel et al., 2010). Specifically, the fact that both the dual eigenvector centrality and the dual Beta-centrality outperform the corresponding metrics computed through the classical approach, indicates that it is important to select partners based on the projects they participate in, especially partners with no previous form of collaboration. Indeed, if an organization joins a new project with another organization that already collaborates with, the centrality of the project increases, but the centrality of the organization remains almost stable. On the contrary, if an organization joins a new project with a new partner, then both the centrality of the project and the centrality of the organization change. This is the main difference between the classical and dual approach, as in the dual-projection method both values of centrality are considered. Therefore, our findings provide new evidence about the importance of increasing the social capital of an organization in collaborative contexts by establishing new partnerships to have a higher probability of obtaining more funds in the field of European research projects. In particular, since the dual Beta-centrality exhibits the highest coefficient and best model fit compared to all other centrality measures, the significance of partners' connections, and consequently, the effect of diversifying partners, diminishes as we move further away from the focal organization. This suggests that the increased knowledge absorbed from a diverse set of actors wears off at a certain point, which will be interesting to investigate in future research. From a managerial side, our findings suggest that in collaborative research projects, particularly EU-funded ones, it is highly valuable to build a broader network of partners, even though this strategy may require greater coordination efforts, compared to relying on higher levels of trust and established routines, as in the case of strategic alliances (Li et al., 2008) and university-industry collaborations (Bruneel et al., 2010). From a policy perspective instead, it is important to highlight that our results denote the effectiveness of the RTD Policy adopted by the European Commission to provide funds to EU member states in order to promote excellent science and trans-disciplinary research through cross-collaborations between diversified actors. This aspect leads us to question the impact of EU policies on organizational strategies, especially the extent to which the effect of partner selection mechanisms is driven by the design of research frameworks, a key aspect that deserves investigation in future research.

Another key finding is related to the relevance of the activity type of the coordinator. To the best of our knowledge, our study provides the first empirical evidence about this aspect, particularly identifying that projects with private companies or public institutions as coordinators have a higher probability of getting more funds compared to projects coordinated by research organizations and HEIs. Even in this case, managerial implications are important, possibly determining the selection of the project coordinator according to these results. Regarding the macro-country instead, our findings are more cautious, since not all coefficients are significant, and they are relatively small in magnitude. For what concerns the diversity of participants, our results are controversial compared to the literature (Gattringer et al., 2017; Wu et al., 2019). Indeed, we find that consortia with high levels of country heterogeneity are more likely to obtain a higher amount of funds, whereas the effect seems to be the opposite if we consider diversity based on activity type. However, we are prudent in commenting on these findings since the number of different activity types is limited in our case, and the actual composition of participants in terms of country and activity type is also driven by the European Commission's specific requests in the calls for projects.

Finally, we acknowledge some possible limitations of our study. First, there may be potential issues of omitted variable biases affecting the model, which could limit the ability of the main independent variables to explain the variance in the dependent variable. Thus, in the future, we aim to expand this research by combining different datasets to control for endogeneity issues, thereby expecting an increase in the adjusted R-squared. Second, we want to analyze other case studies in the field of collaborative research projects to understand if our findings can be generalized beyond the context of EU FPs. Finally, given the relevance of dual Beta-centrality as supported by our results, we are interested in the analysis of different values of λ_1, in order to understand which is the most efficient value to determine the amount of funds in collaborative research projects.

Chapter 3

Uncovering collaborative patterns and transition dynamics in EU Framework Programmes through network modeling

Abstract

Supporting collaborative R&D has become increasingly relevant for policy-makers. Simultaneously, the innovation performance of organizations is enhanced by the participation in collaborative projects, particularly EU-funded ones. In this paper, we innovatively analyze the projects funded by the first eight EU Framework Programmes (FPs) combining elements from Social Network Analysis and statistics through the lens of economic complexity. We explore collaborative patterns and participation dynamics from FP1 to Horizon 2020 (H2020) by computing participants' centrality and assessing the stochasticity of the process, further estimating the probability of transitioning between classes of centrality across consecutive FPs. In doing so, we contribute to the literature by shedding light on the Markovian nature of collaborative patterns through the investigation of the participation dynamics over an extended period. Our results show the existence of "preferential attachment" mechanisms, thus confirming the relevance of participating in EU-funded projects to strengthen organizations' popularity. On the other hand, we highlight the urgency of addressing "oligopolistic" behaviors linked to European funds that hinder the full realization of the European Research Area (ERA). The crucial role of European policies is emphasized by the estimated transition probabilities, which are influenced by breakthrough events in the EU research framework like the Treaty of Maastricht and the promotion of the ERA. However, sustained efforts are necessary to ensure the persistence of a certain degree of openness and the "democratization" of European research funds.

Keywords: Collaborative research projects, Transition dynamics, Process Markovianity, Economic

complexity, Social Network Analysis, EU Framework Programmes.

3.1 Introduction

A systematic European Research and Technological Development (RTD) policy was established in the 1980s alongside the first European Union Framework Programme (EU FP). The RTD policy aligns with the broader cohesion policy, which envisions the EU as a common market promoting the free circulation of people, goods, capital, and knowledge exchange. The main goal of the EU FPs is to provide funds mainly to EU member states, but also to associate and third countries, in order to promote international research collaboration both at the individual and at the organization level.

Supporting collaborative R&D projects has become increasingly relevant for policy-makers and institutions. Funding is a crucial means to enhance research productivity, even more than collaboration itself. The effects of collaboration indeed, become particularly relevant in the post-funding period, suggesting that it takes time to develop effective collaboration structures that have an impact on research productivity (Defazio et al., 2009). At the European level, collaborations between knowledge-intensive and lagging-behind regions positively affect the innovation capability of the latter ones, demonstrating that a greater openness in the collaboration networks fosters knowledge exchange and spillovers (De Noni et al., 2018). As a consequence, knowledge convergence is gradually emerging among NUTS 2 regions over time, creating an opportunity to integrate knowledge cohesion with the social and economic pillars of the EU cohesion policy (Erdil et al., 2022).

At the same time, organizations benefit from collaborative R&D. Structural properties of collaborative networks are in fact crucial for the firms' innovation performance. Specifically, a high clustering enables an efficient information transmission capacity, while a high reach is associated with a greater diversity of the information flow (Schilling and Phelps, 2007). In particular, participation in R&D projects funded by European initiatives positively affects the firm performance, after or during the project completion depending on the nature of the firm (specifically if a manufacturing or a non-manufacturing firm) (Bayona-Sáez and García-Marco, 2010). For this reason, organizations have increasingly embraced the paradigm of external R&D (Enkel et al., 2009), which is central in the logic of open innovation (Chesbrough, 2003).

From an academic perspective, several authors have investigated the phenomenon of collaborative R&D in innovation and policy-related fields. Different forms of collaborative research have been examined, including joint patent publications (Choe and Lee, 2017; Li et al., 2021), scientific collaborations among authors (Chen et al., 2019; Whetsell, 2023), and collaborative research projects (Chen et al., 2020; Zhang and Chen, 2022). In particular, within the latter research strand, the analysis of projects funded by the EU FPs is one of the most relevant case studies to be explored both in a scientific and in a policy context.

Various approaches have been applied to the study of collaborative R&D. However, the dynamic nature of socio-economic processes necessitates examining them through the new lens of economic complexity. This approach incorporates techniques from spectral analysis and network theory to reduce the complexity of phenomena while preserving as much information as possible, avoiding excessive aggregation (Balland et al., 2022). Social Network Analysis (SNA) especially, has been increasingly adopted to investigate the behaviors of collaborative relationships (Cerqueti et al., 2023), analyze the structure of innovation systems (Ancona et al., 2023b), and identify the key actors in collaboration networks (Cinelli et al., 2022). However, when considering the dynamic evolution of collaborative networks, a controversial aspect emerges in the literature. Most of the works indeed, employ Stochastic Actor-Oriented Models (SAOMs) to explore the mechanisms driving network evolution (Giuliani, 2013; Cao et al., 2017) since they are efficient in processing longitudinal network data (Broekel et al., 2014). Nevertheless, this method assumes that network structures evolve as a Markov chain outcome (Snijders, 2017), whereas in many real networks, evolution often displays a non-Markovian behavior (Williams et al., 2022). Moreover, previous studies on EU FPs focus mainly on macro-level analyses (i.e., at the country or regional level) and average dynamics of specific FPs. Low attention has been paid to the micro-dynamics at the participant level, especially over an extended period, spanning multiple FPs.

In this paper, we aim to tackle both the aforementioned issues by exploring the participation dynamics in collaborative research projects funded by the first eight EU FPs, i.e., from FP1 to Horizon 2020 (H2020). Specifically, we map the local behaviors of single actors in terms of their position in the collaborative networks through the use of centrality measures. Then, we statistically assess whether the dynamics of collaborations among the organizations receiving funds in all the first eight EU FPs have a Markovian nature. Finally, we employ an innovative method to partition the rankings of organizations based on the values of strength centrality, and we estimate the probability of moving from one class of centrality to another across consecutive FPs. Through this approach, we can provide valuable insights into the evolution of European collaborative research networks, highlighting eventual phenomena of path dependency and "oligopolistic" behaviors in European funds. The analysis of the process Markovianity in particular, aims to offer new perspectives to policy-makers and innovation scholars into the role of prior participation and the tendency of organizations to co-develop R&D projects based on historical trajectories. Furthermore, this study contributes to shedding light on the collaboration patterns that are more likely to guide organizations toward core membership in the long term.

In this way, we aim to expand the strand of literature focused on the analysis of collaborative research networks, specifically in the context of European research projects. To the best of our knowledge, this is the first attempt to take into account all the past EU FPs, from FP1 to H2020. This continuity across consecutive FPs allows us to analyze the participation dynamics in European research projects over a longer time horizon compared to previous studies. Additionally, the statistical analysis of the Markovianity of a collaborative process represents a novelty in the field. Beyond its theoretical contributions, this study has relevant policy implications by offering empirical evidence about the effectiveness of European research policies across different FPs and their impact on the promotion of the European Research Area (ERA). Moreover, characterizing the micro-dynamics at the participant level and estimating the transition probabilities from one class of centrality to another will provide important indications to participant organizations in relation to effective collaboration strategies.

The remainder of the paper is organized as follows. An overview of the state of the art on collaborative research networks and EU FPs is provided in Section 3.2. We introduce the case study and our data source in Section 3.3. Section 3.4 describes the methodological approach, which is grounded on SNA and statistics. Finally, the results are presented and discussed in Section 3.5, while Section 3.6 highlights the main contributions and possible further developments of this work.

3.2 Literature review and background

The theoretical background of collaborative R&D is extensive, encompassing various research streams and case studies. In this paper, we concentrate on two specific areas of interest: the dynamics of collaborative research networks and the analysis of EU-funded projects. In what follows, we provide an overview of the state of the art of both research streams.

3.2.1 Dynamics of collaborative research networks

The position of research institutes in collaboration networks with industries and/or universities affects their scientific performance. Specifically, the exerted effect depends on the configuration of the network and the types of actors (Chen et al., 2020). Zhang and Chen (2022) show that network position and research performance co-evolve in a mutually reinforcing relationship which is further strengthened by the cognitive proximity between organizations within the collaborative network.

By applying SNA techniques, Xie and Su (2021) observe the “core-periphery” structure of an inter-city research collaboration network, particularly emphasizing its decentralizing trend over time. Cao et al. (2017) observe a similar behavior from the analysis of the project-based collaborative network in the building sector, where the evolution is driven by core-periphery structures and preferential attachment mechanisms, as well as ownership similarities between organizations. R&D networks indeed, evolve based on the choice of collaboration partners, which is in turn influenced by cyclic closure preferences and preferential attachment (Hanaki et al., 2010). Besides endogenous network processes, like preferential attachment and transitivity, international research collaborations are positively affected by the level of liberal democracy and political governance similarities between countries (Whetsell, 2023). Additionally, Zirulia (2023) finds that the long-term dynamics of R&D networks are influenced by events occurring in the early stages of the industry as effect of self-reinforcing processes and a certain path-dependency.

A pioneering work in this respect is that of Wagner and Leydesdorff (2005), who provide evidence of the self-organizing nature of scientific collaboration networks mainly driven by mechanisms of preferential attachment. This is an important outcome from a policy perspective since it suggests that the dynamics of scientific collaboration is based on self-organizing and individual features rather than institutional or policy-related factors. On the other hand, Park and Leydesdorff (2010) show that the dynamics of collaborative networks involving Korean universities, industry, and government are highly sensitive to the research policies implemented by the national government. In particular, those policies prioritizing publication performance tend to discourage collaborations among these three types of actors. Similarly, science and technology policies adopted by the Chinese government are able to influence the evolving small-world structure of scientific collaboration networks involving Triple Helix actors (Zhang et al., 2016). Jakobsen et al. (2019) instead, examine how R&D alliances respond to environmental policies. Their analysis reveals that partners with a high level of relative absorptive capacity (i.e., they are similar in terms of organizational structures, knowledge bases, and dominant logics) are more adept at aligning with policy objectives, especially when there is a balance of power between the two parties, and both dominant logics are considered.

Palla et al. (2007) provide valuable insights into the differences between the dynamics of small collaborative groups and those of large communities. The stability of the formers is guaranteed by the persistence of a few strong ties, whereas large groups rely on continuous changes and member replacements. Moreover, collaborations among organizations belonging to rich clubs (i.e., cohesive groups of core actors in the network) have a particularly high impact on research grants (Nakajima et al., 2023).

The collaborative behaviors of firms typically depend on the technological dynamism of the industry they belong. In technologically dynamic sectors, firms tend to exhibit a higher degree of openness, whereas in stable industries, firms often rely on existing connections and resources (Tatarynowicz et al., 2016). A mechanism of intertemporal diversification of partner types is also highlighted in (Belderbos et al., 2018), where authors demonstrate a strong relationship between the establishment of new collaborative R&D ties and prior collaborations at different value chain levels.

Finally, interesting applications of SNA can be found in the domain of joint patent publications.

Li et al. (2021) investigate the evolution of the patent collaboration network of a Chinese industrial sector. In particular, they focus on the temporal analysis of the overall network structure and individual level characteristics, taking into account the spatial dimension of connections. A similar approach is employed by Choe and Lee (2017) to investigate the Korean patent collaboration networks over time. Their study highlights the small-word structure of the collaborative networks and the increasing key role of universities. However, both studies do not aim to model the dynamics of the whole process, but analyze the changes in the network structure in a more descriptive way.

3.2.2 EU-funded projects

The analysis of EU-funded projects has gained particular momentum in the recent past, stemming from the papers of Breschi and Cusmano (2006) and Roediger-Schluga and Barber (2008), which are among the first studies to explore the structure of collaborative research networks funded by the EU FPs.

Participating in EU research projects is particularly relevant for organizations, significantly enhancing their scientific reputation (Calignano, 2021). Moreover, subsidies received by the EU FPs show a spillover effect on the private dimension of companies, stimulating internal R&D investments, especially for small firms and firms participating in low-budget projects. Overall, participation in projects funded by Horizon 2020 has led to an increase in firms' own R&D spending, while previous FPs appear to be less efficient in achieving this outcome (Szücs, 2020). To this aim, venture capital (VC)-backing plays a crucial role in promoting the participation of new technology-based firms in EU-funded projects, particularly when firms have no or limited experience in this kind of R&D collaborations (Colombo et al., 2016).

On the other hand, the likelihood of receiving funding is influenced by the scientific reputation of applicants (Enger and Castellacci, 2016). This tendency leads the European collaborative research networks to assume oligarchic behaviors, where groups of organizations with high reputation become core actors (Breschi and Cusmano, 2006; Heller-Schuh et al., 2011). In this vein, Lepori et al. (2015) explores the participation of higher education institutions (HEIs) in projects funded by the EU FPs, confirming a direct effect of HEI reputation on their collaboration structure.

In addition to scientific reputation, the centrality of organizations (i.e., their relevance based on their connection structure) in collaborative networks positively affects the likelihood of applying to a EU FP, as well as the probability of obtaining funds (Enger, 2018). Building on these findings, Cinelli et al. (2022) demonstrate that organizations particularly central in the network, measured by a new metric called collective network effect (CNE), have access to more funds than those with lower values of CNE. Through a novel application of SNA, Cerqueti et al. (2023) conduct a rank-size analysis of organizations participating in FP7 and H2020, demonstrating that the universal law well fits the dynamics of the rank-size curves over time. Generally, participants from EU-15 countries are higher ranked in terms of their centrality than those from EU-13, associated, and third countries. Moreover, HEIs tend to exhibit higher centrality values than private companies (Balland et al., 2019).

The success of EU-funded projects is influenced by consortium characteristics, including the experience of the coordinator (Wanzenböck et al., 2020) and geographical heterogeneity among partners (de Arroyabe et al., 2021). Thus, consortium building mechanisms are extremely relevant in EU-funded projects. This underscores the importance of properly considering institutional barries, which could potentially undermine efficient consortium building mechanisms (Cavallaro and Lepori, 2021). Additionally, the efficient design of process, structure, and governance subsystems boosts the performance of EU-funded collaboration networks (Arranz and de Arroyabe, 2012).

Finally, several authors focus on the factors affecting collaboration patterns in EU-funded projects. Among them, geographical distance between organizations emerges as a key determinant of collaborative relationships, negatively impacting the probability of partnering (Scherngell and Barber, 2011). This is a potential challenge to the realization of the ERA (Chessa et al., 2013), although this effect has decreased in more recent FPs (Scherngell and Lata, 2013). Besides geograpical proximity, factors such as social, organizational, technological and economic relatedness also play a role in determining the likelihood of collaborating between two organizations (Paier and Scherngell, 2011; Heringa et al., 2016; Amoroso et al., 2018). Conversely, co-publishing does not appear to drive joint participation in collaborative projects or significantly influence the amount of received funds (Hoekman et al., 2013).

3.3 Case study and data

One of the most relevant case studies in the field of collaborative R&D is represented by the research projects funded by the EU FPs. An essential role of the EU FPs is to provide funds for transnational networks of researchers in order to foster international research collaboration, in light of promoting the ERA. Launched at the Lisbon European Council in March 2000, the ERA initiative has become the central pillar of EU research activities. The emergence of the ERA represents the attempt to make the EU the world's most competitive and dynamic knowledge-based economy (Commission of the European Communities, 2002). According to the Lisbon agenda, the ERA should satisfy a series of principles: an adequate flow of competent researchers; world-class research infrastructures; excellent research institutions; effective knowledge sharing; well-coordinated research programmes and priorities; a wide openess to the world (Commission of the European Communities, 2007). Moreover, since the sixth FP, research institutions have been encouraged to create “centres of excellence” acting as catalysts for marginal actors to increase cooperation and knowledge exchange. In this vein, one of the main purposes is to promote strong and durable partnerships between public and private organizations from different countries (Commission of the European Communities, 2007).

In order to explore these collaboration dynamics over an extended period, we consider as a case study the analysis of projects funded by the first eight EU FPs, from FP1 to H2020. The list of projects funded by the EU FPs, and the related participant organizations, are publicly available from the Community Research and Development Information Service (CORDIS) website[1]. CORDIS serves as the European Commission’s primary repository for project results from the EU FPs. It offers a unique and structured collection of information covering all projects funded from FP1 to Horizon Europe (i.e., the current FP), including details about the associated participants. However, raw data from CORDIS website (especially data referred to the oldest FPs) contains various sources of errors related to the use of several distinct names to address the same organization. This may cause the misallocation of information associated with the same organization, thus undermining the robustness of the results in a dynamic analysis of the EU FPs (Ancona et al., 2023a).

Roediger-Schluga and Barber (2008) addressed this issue by introducing a novel dataset in which authors disambiguate organization names as a result of a labor-intensive procedure involving consistent manual inspection of participants. This effort resulted in the creation of a higher-quality data source known as “EUPRO database”. Most of the studies reported above, exploring the mechanisms of EU-funded projects, base their analyses on the EUPRO database (Paier and Scherngell, 2011; Scherngell and Barber, 2011; Hoekman et al., 2013; Scherngell and Lata, 2013; Lepori et al., 2015; Heringa et al., 2016; Cavallaro and Lepori, 2021).

Therefore, we requested and obtained the access to the EUPRO database containing the list of projects funded by the first eight EU FPs, i.e., FP1 (1984-1987), FP2 (1987-1991), FP3 (1990-1994), FP4 (1994-1998), FP5 (1998-2002), FP6 (2002-2006), FP7 (2007-2013), and H2020 (2014-2020). While data about the first seven FPs are complete, H2020 is still currently updated by data owners, thus results related to the eighth FP may be slightly biased. The number of distinct organizations and distinct projects included in the EUPRO database is reported in Table 3.1.

We anticipate that, for our investigation, we focus on the sample of organizations that received funds from all the first eight FPs (called “common organizations” henceforth). Specifically, we calculate centrality values for all organizations in the networks. Subsequently, we analyze the Markovianity and estimate the transition dynamics of the common organizations participating throughout the entire process from FP1 to H2020. The amount of common organizations is equal to 509.

[1] https://cordis.europa.eu/

Table 3.1: EUPRO database.

	Number of distinct projects	Number of distinct organizations
FP1	3,266	1,972
FP2	3,972	4,587
FP3	5,461	7,095
FP4	14,493	19,255
FP5	15,091	22,862
FP6	10,100	20,582
FP7	25,778	29,334
H2020	25,604	31,319

3.4 Methodological approach

As introduced in Section 3.1, the analysis of collaborative research projects has been approached in various ways in the literature. In this paper, we adopt an innovative approach to the field which combines elements from SNA and statistics under the new lens of economic complexity. More specifically, we explore collaborative patterns and participation dynamics in projects funded by the first eight EU FPs through a three-step procedure. First, we compute the centrality of participants to map their positions in the collaborative networks, and we introduce a ranking of organizations based on their centrality values. Second, we statistically assess whether the collaborative patterns from FP1 to H2020 evolve as an outcome of a Markov chain, i.e., the evolutionary path is memoryless and it is independent of the previous steps of the process. Third, we identify endogenous partitions of common organizations according to their centralities, and we estimate the probability of moving from one class of centrality to another across consecutive FPs. The methodological approach underlying all three steps is thoroughly explained in the following sections.

3.4.1 Social Network Analysis and collaborative networks

SNA stands out as a powerful tool to conceptualize and explore connections among social entities within the framework of economic complexity. Complex systems can in fact be modeled as networks of interacting entities, distinctively characterized by specific structural features (Barabási, 2016; Newman, 2018; Russell and Smorodinskaya, 2018). In these systems, interactions are typically non-linear and heterogeneous (Thurner et al., 2018). Thus, networks are widely used to examine the interactions among the members of a complex system, and investigate structures, roles and dynamics of such systems (Börner et al., 2007). Specifically, common applications of SNA focuses on the analysis of connection structures (Hu et al., 2015) and measures of centrality (Scott and Carrington, 2011).

SNA is grounded on graph theory. Indeed, a network is represented by graph $G = (V, E)$, where $V = \{1, \ldots, n\}$ is the set of vertices (or nodes), and E is the set of edges (or links). Two nodes i and j are adjacent if $(i, j) \in E$. The adjacency relations between nodes are described by a n-square binary matrix A (where n is the number of nodes of a graph), which is called adjacency matrix. A generic element a_{ij} of A is equal to 1 if $(i, j) \in E$ (i.e., the two nodes are connected), and 0 otherwise. In case of a weighted graph, i.e., there is a weight on edges representing factors such as the frequency of occurrence of the associated relation (or, in general, the intensity of the connection), adjacency relations between nodes are described by a weighted adjacency matrix W with zero diagonal entries, and all off-diagonal elements equal to the weight $w_{ij} > 0$ if $(i, j) \in E$, and 0 otherwise.

Most of the studies presented in Section 3.2 adopt SNA techniques to explore the phenomenon of collaborative research. The concept of collaborative networks is indeed widely recognized in the innovation (Nieto and Santamaría, 2007; Tsai, 2009) and R&D (Campos et al., 2013; König et al., 2019) literatures. Based on the network representation, individual organizations engaged in any form of partnership or collaboration are the network nodes, with edges standing for the collaborative relationship between them.

Data about participation in collaborative projects falls into the broad category of two-mode data. This type of data is characterized by two distinct sets of nodes (e.g., actors and events) connected by a relation (e.g., "actors participate in events"), with no connections among nodes in the same set. In our context, the two sets of nodes represent organizations and projects, linked through the relation "organizations participate in projects". Two-mode data is effectively represented in SNA by a bipartite graph. A graph is considered bipartite if its set of nodes V can be partitioned into two different subsets, V_1 and V_2, such that there are no internal links within V_1 and V_2, and all edges connect a node in V_1 and a node in V_2. According to Borgatti and Halgin (2011), the most popular approach to deal with two-mode data is the so called "conversion" approach. This method involves converting two-mode data into two one-mode projections, with one of them usually being of specific interest.

In this study, we work with one-mode projected networks whose nodes correspond to the distinct participants in projects funded by the first eight EU FPs. Consequently, two nodes are connected if the corresponding organizations are partners in one or more projects within the same FP. Thus, participants in projects funded by a specific FP are ultimately represented by an undirected, weighted network, where weights on edges signify multiple partnerships between the same organizations in different projects. Therefore, our case study involves the analysis of eight distinct one-mode projected networks, each representing a different FP.

The most effective way to assess the significance of an organization in a collaborative network is by computing its centrality measures. Centrality is indeed one of the fundamental concepts in

SNA, aiming to determine the importance of a node based on both the quantity and quality of its connections. In SNA, various centrality measures have gained popularity. One of the most intuitive measures is degree centrality, which is equal to the number of adjacent nodes to a certain vertex. Given that our analyses are conducted on weighted networks, we opt to use a variant of degree centrality, specifically the strength centrality, as a proxy for the relevance of participant organizations. Formally, the strength centrality (called "strength" hereafter) of a node i is defined as:

$$s_i = \sum_{j=1}^{n} w_{ij} \tag{3.1}$$

Thus, the greater the number of projects an organization is involved in, the more significant its role within the collaborative network, regardless of common partners across different projects.

Therefore, we calculate a vector of centralities for each FP, where the components correspond to the value of strength of the common organizations receiving funds from FP1 to H2020. Subsequently, we rank the organizations from the highest to the lowest strength values.

3.4.2 Statistical analysis of the process Markovianity

The motivation for evaluating the Markovian nature of participation dynamics in projects funded by the EU FPs arises from the existing literature. In fact, there is a misalignment between empirical studies delving into the evolution of collaborative networks and the theoretical background on network dynamics. In the former case, the vast majority of authors adopt Stochastic Actor-Oriented Models (SAOMs) to dynamically examine collaborative network patterns (Giuliani, 2013; Cao et al., 2017). While these models are efficient in processing longitudinal network data (Broekel et al., 2014), they rely on various assumptions. One of the assumptions is that the evolution of the structure of the analyzed collaborative network is the result of a Markov process (Snijders, 2017). Nonetheless, according to theory, many real networks exhibit a non-Markovian nature, and evolve through a phenomenon of autocorrelation of connections and dynamic correlation among neighboring links (Williams et al., 2022).

In this paper, we aim to address this gap by statistically investigating the Markovianity of the transition dynamics from FP1 to H2020. First, it is worth recalling that a first-order Markov process is characterized by memorylessness, wherein the probability of transitioning to one of the states in the chain in the next step depends solely on the current state, while it is independent of the past states of the process (Gudivada et al., 2015). Such a property has relevant implications in the context of forecasting.

Given a discrete-time stochastic process $X = (X(t) : t \in \mathbb{N})$ assuming values in a set R of rankings (reflecting the strength of organizations in descending order), and denoting as P the related probability law, we can say that X is a Markov chain of order one if it satisfies the following property:

$$P(X(t+1) = i_{t+1}|X(t) = i_t) = P(X(t+1) = i_{t+1}|X(t) = i_t, \ldots, X(1) = i_1, X(0) = i_0), \quad (3.2)$$

for each $t \in \mathbb{N}$, and $i_0, i_1, \ldots, i_t, i_{t+1} \in R$.

The estimation of the second term in Eq. (3.2) is notably complex from a computational perspective. However, according to Renner et al. (2001) and Friedrich et al. (2011), the first-order Markovianity of the process X can be verified by simplifying the condition above as follows:

$$P(X(t+1) = i_{t+1}|X(t) = i_t) = P(X(t+1) = i_{t+1}|X(t) = i_t, X(t-1) = i_{t-1}), \quad (3.3)$$

for each $t \in \mathbb{N}$, and $i_{t-1}, i_t, i_{t+1} \in R$.

In essence, the Markovianity of the process is assessed by comparing the empirical first-order (on the left-hand side of Eq. (3.3)) and second-order (on the right-hand side of Eq. (3.3)) transition matrices.

We adopt an approach similar to that implemented by Cerqueti et al. (2022) for the analysis of ranked textual data. In particular, we start from an empirical sample of $4,072$ observations representing the rankings associated with the strength values of the common organizations across all eight FPs. Subsequently, we extract the sets of one-step and two-step transitions, resulting in $3,563$ pairs of rankings and $3,054$ triples of rankings, respectively. Then, we compute the first-order and second-order transition probability matrices. Starting from the empirical first-order and second-order matrices, we can approximate a statistical assessment of the Markovianity of the process, as described above.

More specifically, we generate $1,000$ processes from both the first-order and second-order transition matrices. Then, the simulated series are pairwise compared using the Kolmogorov-Smirnov (KS) test. This statistical test enables the comparison of empirical distributions between each series generated from the first-order transition matrix and those from the second-order transition matrix. If the two sets of generated series are determined to likely originate from the same distribution, then Eq. (3.3) is verified, and the process can be considered Markovian.

3.4.3 Partitioning the transition probability matrices

One of the primary objectives of this paper is to estimate the probability of transitioning from one class of strength to another across consecutive FPs. Such analysis will provide novel insights into the effectiveness of EU research policies in shaping the structure of funded projects over the years, particularly addressing the risk of "oligopolistic" behaviors (i.e., low ranked organizations do not easily become core actors, and their gap with key members tends to increase over time).

For each pair of consecutive FPs, we can rely on the empirical probability matrix P, whose dimension is equal to $l \times m$, where l is the number of distinct values of strength in a specific FP, and m is the number of distinct values of strength in the subsequent FP. For instance, let us consider the empirical probability matrix P from FP1 to FP2, where the rows consist of the vector of distinct values of strength in FP1, denoted as $s = (s_1, \dots, s_l)$ in ascending order, while the columns represent the vector of distinct values of strength in FP2, denoted as $s' = (s'_1, \dots, s'_m)$ in ascending order. Thus, P has the following structure:

$$P = \begin{bmatrix} p_{1,1} & \cdots & p_{1,m} \\ \vdots & \ddots & \vdots \\ p_{l,1} & \cdots & p_{l,m} \end{bmatrix}, \tag{3.4}$$

where the generic element $p_{i,j}$ is equal to the probability that an organization whose strength in FP1 is s_i has a value of strength in FP2 equal to s'_j. In particular, the probability is computed as follows:

$$p_{i,j} = \frac{k_{i,j}}{\sum_{h=1}^{m} k_{i,h}}, \tag{3.5}$$

where $k_{i,j}$ corresponds to the number of times s_i is associated to s'_j.

However, the empirical probability matrix P assigns a probability to every possible combination of strength values in FP1 and FP2. This analysis would provide a level of detail that might be challenging to interpret in a proper way, given the sparsity of the probability matrix P, mainly filled with 0's and 1's. From a policy perspective, it would be much more relevant to partition the centrality values in classes (e.g., low, medium, and high) in order to estimate the probability of moving, for instance, from a low to a high class of centrality. This process would facilitate the exploration of potential path dependencies in collaborative research projects funded by the EU FPs.

To this aim, the conventional approach involves determining partitions exogenously by establishing either fixed or quantile thresholds of the centrality distribution. Nevertheless, this approach introduces biases in the analysis due to the subjective choices of the thresholds. For this reason, Cerqueti et al. (2017) adopt a novel methodology for identifying partitions in time series endogenously. This method aims to minimize the distance between the estimated probability of moving from one class of centrality to another and the empirical transition probability between distinct centrality values. In this paper, we adapt this approach to partition the distinct values of strength in each FP. Specifically, distinct strength values are grouped together based on the likelihood of obtaining partitions that evolve as similarly to the empirical transitions as possible.

In particular, we aim to identify three classes of strength for each FP: low, medium, and high. To do this, we need to determine two distinct thresholds, t_1 and t_2, by solving the optimization problem described below. Specifically, t_1 and t_2 must correspond to two distinct values of strength such that:

$$Low = \{s_i | s_i \leq t_1\}$$
$$Medium = \{s_i | t_1 < s_i \leq t_2\}$$
$$High = \{s_i | t_2 < s_i\}$$

It is important to highlight that we partition row values for each empirical probability matrix describing the transitions across two consecutive FPs. For all possible partitions, we start by estimating a theoretical probability matrix Π based on a natural assumption of the states' random extraction from the classes. Specifically, we define matrix Π as follows:

$$\Pi = \begin{bmatrix} \pi_{1,1} & \cdots & \pi_{1,m} \\ \vdots & \ddots & \vdots \\ \pi_{l,1} & \cdots & \pi_{l,m} \end{bmatrix} \tag{3.6}$$

Now, let us suppose partitioning the set of distinct values of strength (for instance, in FP1) in the following way:

$$Low = \{s_1, \ldots, s_h\}$$
$$Medium = \{s_{h+1}, \ldots, s_k\}$$
$$High = \{s_{k+1}, \ldots, s_l\}$$

Recalling that the distinct values of strength are expressed in ascending order (so that $s_1 < s_h < s_k < s_l$), we would have that:

$$\pi_{1,j} = \ldots = \pi_{h,j} = \frac{\sum_{i=1}^{h} p_{i,j}}{|Low|} = \frac{\sum_{i=1}^{h} p_{i,j}}{h}, (j = 1, \ldots, m) \tag{3.7}$$

$$\pi_{h+1,j} = \ldots = \pi_{k,j} = \frac{\sum_{i=h+1}^{k} p_{i,j}}{|Medium|} = \frac{\sum_{i=h+1}^{k} p_{i,j}}{k-h}, (j = 1, \ldots, m) \tag{3.8}$$

$$\pi_{k+1,j} = \ldots = \pi_{l,j} = \frac{\sum_{i=k+1}^{l} p_{i,j}}{|High|} = \frac{\sum_{i=k+1}^{l} p_{i,j}}{l-k}, (j = 1, \ldots, m) \tag{3.9}$$

Where in Eq. (3.7), Eq. (3.8), and Eq. (3.9), the denominators cannot be equal to zero (i.e., there cannot exist empty classes) due to the condition that t_1 must be different from t_2. Therefore, we select the best partition based on the estimated theoretical probabilities. More precisely, the final partition is the one that minimizes the distance between the empirical probability matrix P and the theoretical probability matrix Π, formally:

$$min(d_{P,\Pi}) = min \sum_{i=1}^{l} \sum_{j=1}^{m} |p_{i,j} - \pi_{i,j}| \tag{3.10}$$

After determining the optimal partition for FP1, we apply the same procedure to the empirical probability matrix describing the transitions from FP2 and FP3, in order to define the appropriate thresholds for FP2 as well. In this way, we derive the final transition probability matrix $\Pi^{1,2}$ from FP1 to FP2, that is a 3×3 squared matrix defined as follows:

$$\Pi^{1,2} = \begin{bmatrix} \pi_{1,1}^{1,2} & \pi_{1,2}^{1,2} & \pi_{1,3}^{1,2} \\ \pi_{2,1}^{1,2} & \pi_{2,2}^{1,2} & \pi_{2,3}^{1,2} \\ \pi_{3,1}^{1,2} & \pi_{3,2}^{1,2} & \pi_{3,3}^{1,2} \end{bmatrix}, \tag{3.11}$$

where:

$$\begin{aligned}
\pi_{1,1}^{1,2} &= \sum_{j=1}^{h'} \pi_{1,j}; \quad \pi_{1,2}^{1,2} = \sum_{j=h'+1}^{k'} \pi_{1,j}; \quad \pi_{1,3}^{1,2} = \sum_{j=k'+1}^{m} \pi_{1,j} \\
\pi_{2,1}^{1,2} &= \sum_{j=1}^{h'} \pi_{h+1,j}; \quad \pi_{2,2}^{1,2} = \sum_{j=h'+1}^{k'} \pi_{h+1,j}; \quad \pi_{2,3}^{1,2} = \sum_{j=k'+1}^{m} \pi_{h+1,j} \\
\pi_{3,1}^{1,2} &= \sum_{j=1}^{h'} \pi_{k+1,j}; \quad \pi_{3,2}^{1,2} = \sum_{j=h'+1}^{k'} \pi_{k+1,j}; \quad \pi_{3,3}^{1,2} = \sum_{j=k'+1}^{m} \pi_{k+1,j}
\end{aligned} \tag{3.12}$$

In the equations above, h' and k' represent the first and second thresholds of FP2, respectively. Here, the interpretation is as follows. The average probability of transitioning from a low class of strength in FP1 to a low class of strength in FP2 is $\pi_{1,1}^{1,2}$. The average probability of transitioning from a low class of strength in FP1 to a medium class of strength in FP2 is $\pi_{1,2}^{1,2}$. The average probability of transitioning from a low class of strength in FP1 to a high class of strength in FP2 is $\pi_{1,3}^{1,2}$, and so on.

The transition probability matrix is computed for all pairs of consecutive FPs. It is important to note that, since H2020 is the last FP included in the EUPRO database, it cannot be partitioned according to the methodology introduced before. Thus, the last transition probability matrix presented in the following Section concerns the transition between FP6 and FP7.

3.5 Results and discussion

The programming budget established by the European Commission to fund the EU FPs has increased over the years. Alongside the allocated funds, there has been an evolution from a European policy perspective, leading to an increasing interconnectedness among European regions. As a consequence, collaborative research projects funded by the EU FPs have been designed to involve a growing number of participant organizations, fostering collaboration among institutions from different countries and sectors.

This tendency has immediately emerged from the data reported in Table 3.1. Indeed, both the number of funded projects and the number of distinct participants steadily increase over time (except from FP5 to FP6). However, what proves particularly interesting to explore is the variation in the distribution of connections, reflecting the participation in projects. To this aim, Figure 3.1 illustrates the trend of mean strength (i.e., the average strength among network members) over the different FPs. In particular, we distinguish between the set of common organizations (i.e., those receiving funding by all the first eight FPs) and the entire collaborative network for each FP.

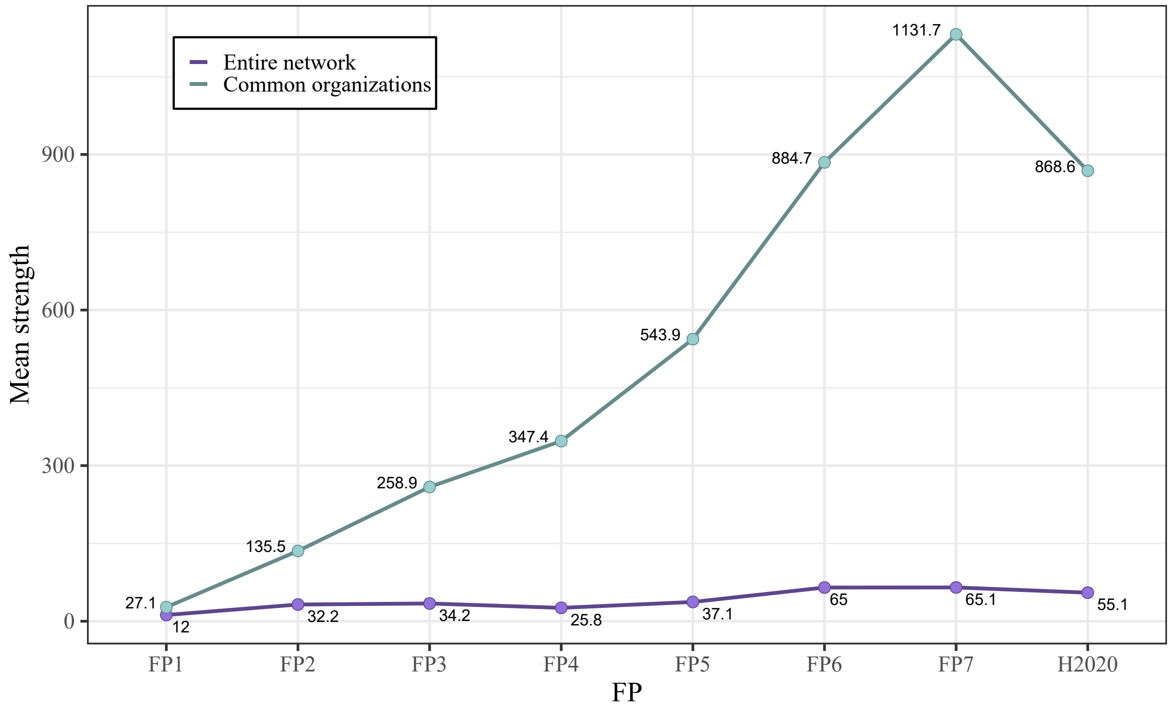


Figure 3.1: Trend of mean strength over different FPs: comparison between the entire network and the common organizations.

This representation offers additional insights into the aforementioned dynamics. What emerges as particularly noteworthy is the atypical outcome observed in relation to H2020. While participations in H2020 are currently updated, thus influencing the results concerning the eighth FP, there may be additional explanations to consider. As shown in Table 3.1, the decrease in the number of funded projects compared to FP7, coupled with an increase in the number of participants across the two FPs, implies a smaller average number of projects per organization in H2020, leading to a declining trend in mean strength. Nonetheless, considering that H2020 was the EU FP with the highest amount of funds prior to Horizon Europe, we anticipate that this result may be influenced by the actual number of projects in the dataset.

Except for H2020, the trend of mean strength exhibits a slight increase from FP1 to FP6 within the entire network (except for a decrease in correspondence with FP4) before becoming stable from FP6 to FP7. This trend aligns with the growing amount of funds provided by the European Commission and the tendency to form consortia including an increasing number of participants in recent FPs compared to the earlier ones. What is interesting to notice however, is the significant increase in the trend of mean strength over the FPs for the common organizations. This suggests that new incumbents in European research projects tend to collaborate with experienced organizations, benefiting from the expertise developed through past participation. However, such a mechanism widens the gap between old and new participants over time, as evident from Figure 3.1. This phenomenon is well known in network theory under the name of "preferential attachment" (Newman, 2001), and it has been observed by several authors in the dynamic analysis of collaborative research networks (Wagner and Leydesdorff, 2005; Hanaki et al., 2010; Cao et al., 2017). Here the implications are twofold. From an organizational perspective, participating in EU-funded projects provides a competitive advantage for future FPs due to growing experience, competencies, and popularity, as well as other network benefits. This finding aligns with previous works emphasizing the advantages of engaging in collaborative research projects, particularly in relation to EU-funded ones. Nevertheless, these kinds of "oligopolistic" behaviors pose a challenge to EU research policies, exacerbating inequalities in the distribution of funds among participants. Despite being a structural property of network evolution, preferential attachment mechanisms should be better addressed by offering incentives or proposing apposite funding schemes to encourage partnerships among newcomers, potentially with the coordinating role of an experienced institution.

3.5.1 Unveiling dynamic patterns of EU Framework Programmes

As described in Section 3.4.2, we evaluate the Markovian nature of participation dynamics in EU FPs by comparing first-order and second-order transition matrices. In more detail, we generate $1,000$

processes from both empirical matrices, reproducing at each step a vector of rankings associated with a hypothetical FP. Subsequently, we conduct pairwise comparisons between the empirical distributions of the series generated from the first-order transition matrix and those from the second-order transition matrix through the KS test. In order to ensure more robust results, we simulate 100 steps for each generated series, allowing us to rely on a distribution coming from a larger number of observations. Even though the observed phenomenon is characterized by an 8-step process, this procedure does not compromise the reliability of the results. Indeed, the probability distribution function from which the 100-step series are generated remains the empirical one, projected over a longer time horizon for a more comprehensive analysis.

The distribution of the $1,000$ values of KS statistics is presented in Figure 3.2, along with the thresholds of the statistics at three different confidence levels (90%, 95%, and 99%). Two representantions are displayed as we employ two distinct methods for generating the $1,000$ series from the second-order transition matrix.

First, it is possible to notice that the two representations are similar. This aspects reveals that the analysis is not sensitive to the method used for generating the series from the second-order transition matrix, thereby strengthening the robustness of our findings. Nevertheless, the results of the KS test are not obvious. In particular, taking into account the highest confidence level (i.e., 99%), we observe a portion of KS statistics values to the left of the threshold, whereas the remaining part of the distribution of the KS statistics does not support the hypothesis of process Markovianity. However, considering that the mode of the distribution falls within the 99% confidence interval, we can conclude that the most likely scenario, in the comparison between first-order and second-order transition matrices, suggests that the process is generated from a Markov chain.

We can then observe a quasi-Markovian nature of the participation dynamics in collaborative research projects funded by the EU FPs. This finding offers novel insights into the evolution of collaborative networks, particularly in the field of EU-funded projects. Previous studies have traditionally treated the evolution of collaborative networks as outcomes of Markov chain processes. In this paper, we aim to focus on the phenomenon's complexity and the need to rigorously evaluate the dynamic patterns of such networks. Moreover, we identify a peculiar behavior in EU-funded projects that differentiates them from other network evolution models which are non-Markovian. The quasi-Markovianity of European collaborative research networks offers novel insights to policy-makers and innovation scholars, demonstrating that R&D partnerships often emerge as a result of prior collaborations, and that historical trajectories can be analyzed to understand the potential path of R&D communities. In fact, the quasi-Markovianity of ranking dynamics opens up opportunities for accurate forecasting procedures aimed at estimating the leaders of future FPs and the probability of success of applicants based on their position in the previous FP.

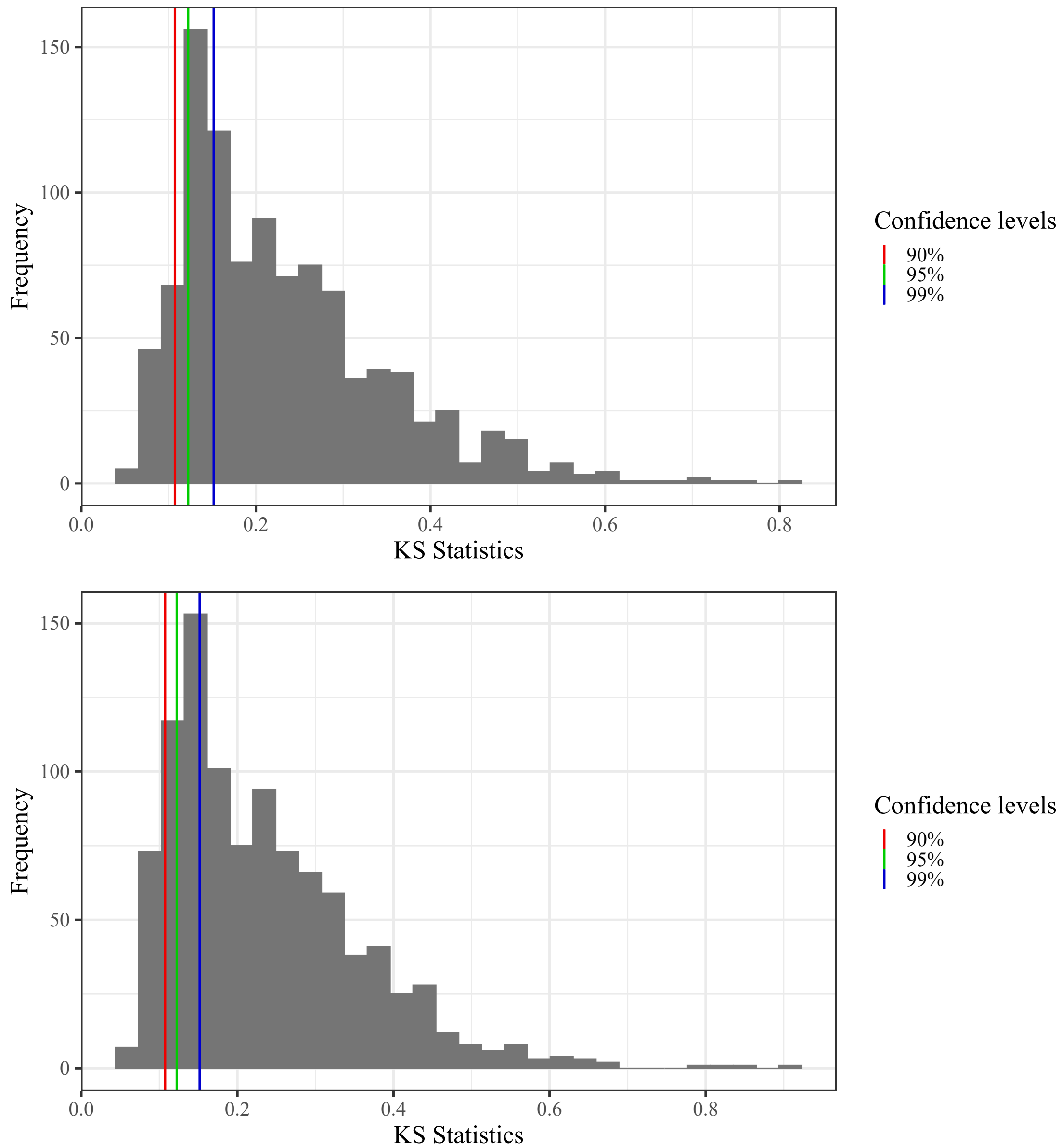


Figure 3.2: Distribution of the KS statistics obtained by comparing the simulated series generated from the first-order and second-order transition matrices. Colored lines represent the thresholds calculated at various confidence levels (90%, 95%, 99%).

3.5.2 The effects of EU initiatives on the transition probability matrices

The final part of this research involves partitioning the transition probability matrices of the common organizations to estimate the probability of moving from one class of strength to another across consecutive FPs. As previously discussed in Section 3.4.3, we identify thresholds to endogenously partition the transition probability matrices into three different classes (i.e., low, medium, and high) of strength values. Thus, for each pair of consecutive FPs, we estimate a 3×3 transition probability matrix.

The steadily increasing trend in the strength values of common organizations, as illustrated in Figure 3.1, leads us to expect a general rise in the class of centrality from a FP to the subsequent one. In other words, transitions from low to medium, low to high, and medium to high are expected to be more likely than transitions from high to low, high to medium, and medium to low. Such expectations are confirmed by the estimated transition probability matrices reported below.

$$\Pi^{1,2} = \begin{bmatrix} 0.36 & 0.14 & 0.50 \\ 0.15 & 0.09 & 0.76 \\ 0.04 & 0.02 & 0.94 \end{bmatrix} ; \quad \Pi^{2,3} = \begin{bmatrix} 0.96 & 0.00 & 0.04 \\ 0.84 & 0.01 & 0.15 \\ 0.24 & 0.01 & 0.75 \end{bmatrix} ; \quad \Pi^{3,4} = \begin{bmatrix} 0.48 & 0.01 & 0.51 \\ 0.00 & 0.00 & 1.00 \\ 0.03 & 0.00 & 0.97 \end{bmatrix}$$

$$\Pi^{4,5} = \begin{bmatrix} 0.33 & 0.02 & 0.65 \\ 0.00 & 0.00 & 1.00 \\ 0.03 & 0.00 & 0.97 \end{bmatrix} ; \quad \Pi^{5,6} = \begin{bmatrix} 0.92 & 0.00 & 0.08 \\ 1.00 & 0.00 & 0.00 \\ 0.28 & 0.01 & 0.71 \end{bmatrix} ; \quad \Pi^{6,7} = \begin{bmatrix} 0.58 & 0.01 & 0.41 \\ 0.17 & 0.00 & 0.83 \\ 0.02 & 0.01 & 0.97 \end{bmatrix}$$

As it is possible to notice indeed, the class of strength of the organizations has a higher probability of increasing rather than decreasing in almost all transition matrices. Additionally, it is rather hard for a participant in a high class of strength to move to a less central position in the subsequent FP. In fact, the probability of remaining a core institution in a high class of strength is always above 0.71, and in four out of six transition matrices, it exceeds 0.90. On the other hand, a participant in a low class of strength is more likely to increase its centrality over consecutive FPs, except for the transitions from FP2 to FP3 and from FP5 to FP6. While the other four transition matrices exhibit a similar structure, $\Pi^{2,3}$ and $\Pi^{5,6}$ are quite peculiar. These are the only two cases indeed, in which the sum of the elements above the main diagonal in both transition matrices is lower than the sum of the elements below the main diagonal. What emerges then, is that organizations are not particularly boosted by their previous participations when moving from FP2 to FP3 and from FP5 to FP6. Diving deep into these dynamics, we may find a connection with significant events that happened during the course of the third and the sixth FPs, respectively. First, the Treaty of Maastricht, which took effect in 1993, changed the legal basis for the deployment of the EU FPs, turning them into financial tools to foster European research activities and opening the research programmes to a wider range of topics. Second, the adoption of the ERA in 2000, but fully effective since FP6, marked a substantial shift in European research activities and funding schemes, as highlighted in Section 3.3. Both events introduced a kind of breakthrough in the functioning of European research programmes and in the allocation of funds, increasing competitiveness and openness while partially diminishing competitive advantages acquired through previous participations.

3.6 Conclusions

In this paper, we employ an innovative approach to the analysis of collaborative research projects funded by the EU FPs, spanning from FP1 to H2020. In particular, we explore collaborative patterns and participation dynamics through the lens of economic complexity. We combine SNA and statistics to map the connection structure of organizations and determine the stochastic nature of their participation process in EU-funded projects. Specifically, we compute the strength centrality (i.e., an extension of degree centrality that considers weights on edges) of organizations receiving funds in all the first eight EU FPs. Subsequently, we conduct a statistical assessment to investigate whether the participation dynamics of these organizations exhibit a Markovian nature through the comparison between the first-order and second-order transition matrices of their rankings. Finally, we partition the values of strength in three classes (i.e., low, medium, and high) and estimate the probability of transitioning from one class of centrality to another across consecutive FPs.

Our findings emphasize the contributions of this study both to theory and practice. From a theoretical point of view, we extend the knowledge on the evolution of collaborative research networks, particularly within the context of the EU FPs. Specifically, we shed light on a controversial aspect in the literature regarding the Markovian nature of collaborative dynamics. The vast majority of authors indeed, implicitly assume collaborative networks evolving as a Markov chain outcome through the implementation of SAOMs to explore their dynamics (Giuliani, 2013; Cao et al., 2017; Snijders, 2017). To the best of our knowledge, this is the first study to systematically examine the stochasticity of the entire participation process in collaborative research networks. Our results offer novel insights into the dynamics of this kind of networks, revealing their quasi-Markovian nature, which distinguishes them from other existing real networks (Williams et al., 2022). One possible motivation for such peculiar behavior is that these mechanisms reflect dynamics induced by the success stories of prior winners of European projects. In particular, obtaining funds during a EU FP under a specific funding scheme makes an organization an attractive partner in that research field for future proposals to similar calls for projects. Also, the results on the first-order Markovianity are consistent with the idea that the previous position is the most critical factor in predicting future roles since the continuity between calls and funding schemes is easier across consecutive FPs. Furthermore, we complement the literature on EU FPs, which has predominantly focused on spatial analyses in a static context. Differently, our study concentrates on the analysis of the micro-dynamics at the participant level over an extended time period, encompassing all the first eight EU FPs. Additionally, from a methodological perspective, this paper contributes to the field by proposing a procedure for endogenously partitioning groups of organizations. This approach enhances the precision and unbiasedness of our results, allowing us to capture the effectiveness of research policies on transition dynamics.

Besides theoretical contributions, our results hold several implications both at the participant and policy levels. The increasing gap between the mean strength of the common organizations and the entire network highlights the existence of "preferential attachment" mechanisms shaping the dynamics of collaborative projects. This finding, consistent with prior studies on the topic (Wagner and Leydesdorff, 2005; Hanaki et al., 2010; Cao et al., 2017), confirms the relevance of participating in EU-funded projects to enhance the scientific reputation of organizations (Calignano, 2021), and strengthen their popularity and competencies. From a policy perspective, our findings underscore the urgency of addressing such "oligopolistic" behaviors associated with European funds, which hinder the full realization of the ERA and maintain the core-periphery structure of collaborative research networks. New consortium-building mechanisms should be promoted, encouraging diversity in experience levels, similar to the principles of country and activity type diversification, which are actually key requisites in EU-funded projects. The crucial role of European policies and interventions on funding scheme mechanisms is further emphasized by our estimation of the probabilities of moving from a class of strength to another across consecutive FPs. Indeed, the transitions from FP2 to FP3 and from FP5 to FP6 exhibit the lowest signal of competitive advantage derived from previous participation in European projects. In correspondence with these transitions, breakthrough events in the EU research framework took place. The Treaty of Maastricht first, and the promotion of the ERA then, marked key milestones, fostering the openness and the "democratization" of European research funds. However, sustained and continuous efforts are required to ensure the persistence of similar effects.

It is important to specify that the estimated transition probabilities are contingent on the dimension of the identified classes. Our approach, which endogenously partitions the distinct values of strength based on their evolving patterns, tends to include a small number of organizations in the medium class compared to the low and high categories. Consequently, values corresponding to medium classes of centralities are more extreme and less significant. Also, the number of observations affects the sparsity of some transition matrices, thus representing a potential limitation of our study. For this reason, in the future, we are interested in extending our analysis to other case studies in the field of collaborative research projects, with the potential benefit of accessing more granular data. Moreover, we aim to develop accurate forecasting procedures for determining the participation in the next EU FPs, leveraging the quasi-Markovian nature of the analyzed process.

Conclusions

This thesis contributes to the literature on entrepreneurial ecosystems and collaborative R&D by adopting an innovative approach to these fields. Specifically, the three chapters revolve around the new paradigm of economic complexity, which is well-suited to analyze the increasing interconnectedness of innovation systems while simultaneously reducing the computational dimension of socio-economic phenomena. In particular, the employment of network theory is central within the thesis, addressing three different issues: (i) the study of the structural features of entrepreneurial ecosystems and the proposal of network metrics for their measurement; (ii) the analysis of the effect of different partner selection mechanisms on the amount of funds in collaborative research projects; (iii) the investigation of participation dynamics in collaborative research projects funded by the EU FPs. The main results and findings reported within each chapter contribute both to the theory (academic knowledge) and practice (managerial and policy implications).

In particular, this thesis contributes to the existing body of literature on entrepreneurial ecosystems and collaborative R&D as follows. Chapter 1 provides insights into the structural features of entrepreneurial ecosystems by identifying seven network-based principles characterizing their structure: connectivity, density, stability, leadership, diversity, intermediaries, and feedback loops. Moreover, the proposition of network metrics to measure the identified structural properties allows this research to contribute methodologically, besides theoretically, filling a gap in the extant literature (Wurth et al., 2022). The proposed approach indeed, allows monitoring the structure of entrepreneurial ecosystems in real-time, avoiding reliance on annually released national and international statistics. Finally, the analyzed case study represents a further contribution to the field by presenting a novel way to conceptualize entrepreneurial ecosystems, considering online interactions among actors. The pandemic has indeed contributed to shifting several relationships toward digital platforms, extending the concept of proximity, which is key in the definition of entrepreneurial ecosystems, beyond geographical boundaries. Chapter 2 extends the current literature that explores the impact of partner selection strategies on the economic contribution received by collaborative research projects. Specifically, this strand of literature lacks evidence on the analysis of strategic choices behind partner selection mechanisms. The main peculiarity of this study is represented by the application of the dual-projection approach (Everett and Borgatti, 2013) to collaborative research networks, through which it is possible to assess the centrality of projects while accounting for the centrality of participant organizations. By comparing classical and dual centrality measures, this study demonstrates the importance of selecting partners based on the projects they participate in, especially partners with no previous form of collaboration. These findings provide new evidence regarding the relevance of increasing the social capital of an organization (Seo, 2020) by establishing new partnerships to have a higher probability of obtaining more funds in the field of European research projects. This insight contributes to differentiating collaborations and strategic alliances, suggesting that in strategic alliances, trust and established routines are more important than increasing social capital, whereas in collaborative research projects, it is highly valuable to build a broader network of partners, despite this strategy may require greater coordination efforts. Finally, Chapter 3 contributes to the literature on the dynamics of collaborative research projects, particularly EU-funded ones, by shedding light on the Markovian nature of participation dynamics, which is commonly assumed by studies employing SAOMs methods for network evolution. In particular, the obtained results reveal the quasi-Markovian nature of such processes, distinguishing them from other existing real networks (Williams et al., 2022). Furthermore, this study expands the literature on EU FPs, which has predominantly focused on spatial analyses in static contexts, by examining micro-dynamics at the participant level across all the first eight EU FPs. Additionally, from a methodological perspective, this chapter introduces a novel procedure for endogenously partitioning groups of organizations based on their evolutionary patterns.

From a policy perspective, Chapter 1 provides a set of metrics to support the measurement of entrepreneurial ecosystems and the development of data-driven policies. The proposed methodology indeed, is designed to offer real-time indications of how an entrepreneurial ecosystem is structurally developing, allowing for timely interventions in case of unmet properties. The identified metrics can assist policy-makers and institutions in funding the establishment of collaborative networks and entrepreneurial and innovation ecosystems, further contributing to evaluating the impact of their policies on the structure of such systems. Thus, this work represents a first attempt to analyze the social network structure of an innovation system to guide policy-making decisions. Chapter 2 offers novel insights into the effectiveness of the EU Research and Technological Development Policy, suggesting that cross-collaboration between diversified actors is positively related to higher economic contributions, in line with the promotion of excellent science. However, at the same time, EU research policies should address the risk of "oligopolistic" behaviors associated with European funds that can potentially hinder the full realization of the ERA. In fact, as highlighted in Chapter 3, the existence of "preferential attachment" mechanisms shapes the dynamics of collaborative projects in EU FPs. To this aim, new consortium-building mechanisms should be promoted, fostering diversity in experience levels, similar to the current practice of country and activity type diversification. The crucial role of European policies influencing the funding schemes of EU research frameworks is also demonstrated by the estimation of the probabilities of moving from one class of strength to another across consecutive FPs. In fact, these transitions are affected by breakthrough events, such as the Treaty of Maastricht and the promotion of the ERA, which represent key initiatives to foster the openness and "democratization" of European research funds. Therefore, sustained and continuous efforts are necessary to ensure the persistence of "anti-oligopolistic" effects.

Finally, managerial implications can be derived from Chapter 2 and Chapter 3 in relation to organizations participating in collaborative projects funded by the EU FPs. Firstly, organizations should consider partnering with previously unrelated actors to increase the likelihood of obtaining more funds, and possibly setting private companies or public institutions as project coordinators. Secondly, results from Chapter 3 provide new insights into the relevance of participating in EU-funded projects to enhance the scientific reputation of organizations (Calignano, 2021) and strengthen their popularity and competencies, aligning with prior studies on the topic (Wagner and Leydesdorff, 2005; Hanaki et al., 2010; Cao et al., 2017).

Bibliography

Acs, Z. J., Stam, E., Audretsch, D. B., and O'Connor, A. (2017). The lineages of the entrepreneurial ecosystem approach. *Small Business Economics*, 49:1–10. https://doi.org/10.1007/s11187-017-9864-8.

Al-Khalifa, A. K. and Eggert Peterson, S. (1999). The partner selection process in international joint ventures. *European Journal of Marketing*, 33(11/12):1064–1081. https://doi.org/10.1108/03090569910292276.

Albert, R., Jeong, H., and Barabási, A. L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382. https://doi.org/10.1038/35019019.

Aleisa, E. (2013). Startup Ecosystems: Study of the ecosystems around the world; Focusing on Silicon Valley, Toronto and Moscow.

Almeida, F., Duarte Santos, J., and Augusto Monteiro, J. (2020). The challenges and opportunities in the digitalization of companies in a post-covid-19 world. *IEEE Engineering Management Review*, 48(3):97–103. https://doi.org/10.1109/EMR.2020.3013206.

Alonso, J. M. and Andrews, R. (2019). Governance by targets and the performance of cross-sector partnerships: Do partner diversity and partnership capabilities matter? *Strategic Management Journal*, 40(4):556–579. https://doi.org/10.1002/smj.2959.

Amoroso, S., Coad, A., and Grassano, N. (2018). European R&D networks: a snapshot from the 7th EU Framework Programme. *Economics of Innovation and New Technology*, 27(5-6):404–419. https://doi.org/10.1080/10438599.2017.1374037.

Ancona, A., Cerqueti, R., and Vagnani, G. (2023a). A novel methodology to disambiguate organization names: an application to EU Framework Programmes data. *Scientometrics*, 128:4447–4474. https://doi.org/10.1007/s11192-023-04746-x.

Ancona, A., Cinelli, M., Ferraro, G., and Iovanella, A. (2023b). Network-based principles of entrepreneurial ecosystems: a case study of a start-up network. *Small Business Economics*, 61:1497–1514. https://doi.org/10.1007/s11187-023-00738-6.

Antikainen, M., Mäkipää, M., and Ahonen, M. (2010). Motivating and supporting collaboration in open innovation. *European Journal of Innovation Management*, 13(1):100–119. https://doi.org/10.1108/14601061011013258.

Arranz, N. and de Arroyabe, J. C. F. (2012). Can innovation network projects result in efficient performance? *Technological Forecasting and Social Change*, 79(3):485–497. https://doi.org/10.1016/j.techfore.2011.09.003.

Atta-Owusu, K., Fitjar, R. D., and Rodríguez-Pose, A. (2021). What drives university-industry collaboration? Research excellence or firm collaboration strategy? *Technological Forecasting and Social Change*, 173:121084. https://doi.org/10.1016/j.techfore.2021.121084.

Audretsch, D. B., Colombelli, A., Grilli, L., Minola, T., and Rasmussen, E. (2020). Innovative start-ups and policy initiatives. *Research Policy*, 49(10):104027. https://doi.org/10.1016/J.RESPOL.2020.104027.

Audretsch, D. B., Cunningham, J. A., Kuratko, D. F., Lehmann, E. E., and Menter, M. (2019). Entrepreneurial ecosystems: economic, technological, and societal impacts. *The Journal of Technology Transfer*, 44:313–325. https://doi.org/10.1007/s10961-018-9690-4.

Auerswald, P. E. and Dani, L. (2017). The adaptive life cycle of entrepreneurial ecosystems: the biotechnology cluster. *Small Business Economics*, 49:97–117. https://doi.org/10.1007/s11187-017-9869-3.

Autio, E., Kenney, M., Mustar, P., Siegel, D., and Wright, M. (2014). Entrepreneurial innovation: The importance of context. *Research Policy*, 43(7):1097–1108. https://doi.org/10.1016/j.respol.2014.01.015.

Autio, E., László, S., Komlósi, E., and Mónika, T. (2020). Eides 2020 - the european index of digital entrepreneurship systems. *Publications Office of the European Union, Luxembourg.* https://doi.org/10.2760/150797.

Autio, E., Nambisan, S., Thomas, L. D., and Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, 12(1):72–95. https://doi.org/10.1002/sej.1266.

Balland, P.-A., Boschma, R., and 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., Broekel, T., Diodato, D., Giuliani, E., Hausmann, R., O'Clery, N., and Rigby, D. (2022). Reprint of the new paradigm of economic complexity. *Research Policy*, 51(8):104568. https://doi.org/10.1016/j.respol.2022.104568.

Barabási, A.-L. (2016). *Network Science*. Cambridge University Press.

Baum, J. A., Cowan, R., and Jonard, N. (2010). Network-independent partner selection and the evolution of innovation networks. *Management Science*, 56(11):2094–2110. https://doi.org/10.1287/mnsc.1100.1229.

Bayona-Sáez, C. and García-Marco, T. (2010). Assessing the effectiveness of the eureka program. *Research Policy*, 39(10):1375–1386. https://doi.org/10.1016/j.respol.2010.07.007.

Belderbos, R., Gilsing, V., Lokshin, B., Carree, M., and Sastre, J. F. (2018). The antecedents of new r&d collaborations with different partner types: On the dynamics of past r&d collaboration and innovative performance. *Long Range Planning*, 51(2):285–302. https://doi.org/10.1016/j.lrp.2017.10.002.

Bengtsson, L., Lakemond, N., Lazzarotti, V., Manzini, R., Pellegrini, L., and Tell, F. (2015). Open to a select few? Matching partners and knowledge content for open innovation performance. *Creativity and Innovation Management*, 24(1):72–86. https://doi.org/10.1111/caim.12098.

Bhullar, S. S., Nangia, V. K., Batish, A., et al. (2019). The impact of academia-industry collaboration on core academic activities: Assessing the latent dimensions. *Technological Forecasting and Social Change*, 145(C):1–11. https://doi.org/10.1016/j.techfore.2019.04.021.

Blank, S. (2010). Why Startups are Agile and Opportunistic – Pivoting the Business Model. *Steve Blank*.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*, volume 7. Free Press New York.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120. https://doi.org/10.1080/0022250X.1972.9989806.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5):1170–1182. https://doi.org/10.1086/228631.

Bonacich, P. (1991). Simultaneous group and individual centralities. *Social Networks*, 13(2):155–168. https://doi.org/10.1016/0378-8733(91)90018-O.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, 29(4):555–564. https://doi.org/10.1016/j.socnet.2007.04.002.

Borgatti, S. P. and Halgin, D. S. (2011). Analyzing affiliation networks. In *Carrington P, Scott J (Eds.), The Sage handbook of social network analysis*. SAGE publications.

Börner, K., Sanyal, S., and Vespignani, A. (2007). Network science. *Annual Review of Information Science and Technology*, 41(1):537–607. https://doi.org/10.1002/aris.2007.1440410119.

Breschi, S. and Cusmano, L. (2006). Unveiling the texture of a European Research Area: emergence of oligarchic networks under EU Framework Programmes. In *Knowledge flows in European industry*, pages 294–324. Routledge.

Broekel, T., Balland, P.-A., Burger, M., and 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.

Bruneel, J., d'Este, P., and 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.

Calignano, G. (2021). Better connected, more reputable? On the association between node centrality and academic reputation in the European Union research and innovation networks. *European Policy Analysis*, 7(1):240–262. https://doi.org/10.1002/epa2.1079.

Campos, P., Brazdil, P., and Mota, I. (2013). Comparing Strategies of Collaborative Networks for R&D: an agent-based study. *Computational Economics*, 42(1):1–22. https://doi.org/10.1007/S10614-013-9376-9.

Cao, D., Li, H., Wang, G., Luo, X., Yang, X., and Tan, D. (2017). Dynamics of project-based collaborative networks for bim implementation: Analysis based on stochastic actor-oriented models. *Journal of Management in Engineering*, 33(3):04016055. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000503.

Cao, Z. and Shi, X. (2021). A systematic literature review of entrepreneurial ecosystems in advanced and emerging economies. *Small Business Economics*, 57:75–110. https://doi.org/10.1007/s11187-020-00326-y.

Carayannis, E. G., Grigoroudis, E., and Wurth, B. (2022). Or for entrepreneurial ecosystems: A problem-oriented review and agenda. *European Journal of Operational Research*, 300(3):791–808. `https://doi.org/10.1016/j.ejor.2021.10.030`.

Cavallaro, M. and Lepori, B. (2021). Institutional barriers to participation in EU framework programs: contrasting the Swiss and UK cases. *Scientometrics*, 126(2):1311–1328. `https://doi.org/10.1007/s11192-020-03810-0`.

Cavallo, A., Ghezzi, A., and Balocco, R. (2019). Entrepreneurial ecosystem research: present debates and future directions. *International Entrepreneurship and Management Journal*, 15(4):1291–1321. `https://doi.org/10.1007/S11365-018-0526-3`.

Cavallo, A., Ghezzi, A., and Sanasi, S. (2021). Assessing entrepreneurial ecosystems through a strategic value network approach: evidence from the San Francisco Area. *Journal of Small Business and Enterprise Development*, 28(2):261–276. `https://doi.org/10.1108/JSBED-05-2019-0148`.

Cerqueti, R., Falbo, P., and Pelizzari, C. (2017). Relevant states and memory in markov chain bootstrapping and simulation. *European Journal of Operational Research*, 256(1):163–177. `https://doi.org/10.1016/j.ejor.2016.06.006`.

Cerqueti, R., Ferraro, G., and Iovanella, A. (2018). A new measure for community structures through indirect social connections. *Expert Systems with Applications*, 114:196–209. `https://doi.org/10.1016/J.ESWA.2018.07.040`.

Cerqueti, R., Ficcadenti, V., Dhesi, G., and Ausloos, M. (2022). Markov chain monte carlo for generating ranked textual data. *Information Sciences*, 610:425–439. `https://doi.org/10.1016/j.ins.2022.07.137`.

Cerqueti, R., Iovanella, A., and Mattera, R. (2023). Clustering networked funded european research activities through rank-size laws. *Annals of Operations Research*, pages 1–29. `https://doi.org/10.1007/s10479-023-05321-6`.

Chan, Y. E. and Farrington, C. J. (2018). Community-based research: Engaging universities in technology-related knowledge exchanges. *Information and Organization*, 28(3):129–139. `https://doi.org/10.1016/j.infoandorg.2018.08.001`.

Chen, K., Zhang, Y., and Fu, X. (2019). International research collaboration: An emerging domain of innovation studies? *Research Policy*, 48(1):149–168. `https://doi.org/10.1016/j.respol.2018.08.005`.

Chen, K., Zhang, Y., Zhu, G., and Mu, R. (2020). Do research institutes benefit from their network positions in research collaboration networks with industries or/and universities? *Technovation*, 94:102002. `https://doi.org/10.1016/j.technovation.2017.10.005`.

Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Chessa, A., Morescalchi, A., Pammolli, F., Penner, O., Petersen, A. M., and Riccaboni, M. (2013). Is Europe evolving toward an integrated research area? *Science*, 339(6120):650–651. `https://doi.org/10.1126/science.1227970`.

Choe, H. and Lee, D. H. (2017). The structure and change of the research collaboration network in korea (2000–2011): Network analysis of joint patents. *Scientometrics*, 111:917–939. `https://doi.org/10.1007/s11192-017-2321-2`.

Cinelli, M., Ferraro, G., and Iovanella, A. (2022). Connections matter: a proxy measure for evaluating network membership with an application to the Seventh Research Framework Programme. *Scientometrics*, 127(7):3959–3976. `https://doi.org/10.1007/s11192-022-04414-6`.

Cinelli, M., Peel, L., Iovanella, A., and Delvenne, J.-C. (2020). Network constraints on the mixing patterns of binary node metadata. *Physical Review E*, 102(6):062310. `https://doi.org/10.1103/PhysRevE.102.062310`.

Clemente, G. P., Fattore, M., and Grassi, R. (2018). Structural comparisons of networks and model-based detection of small-worldness. *Journal of Economic Interaction and Coordination*, 13:117–141. `https://doi.org/10.1007/s11403-017-0202-7`.

Cloutier, L. and Messeghem, K. (2022). Whirlwind model of entrepreneurial ecosystem path dependence. *Small Business Economics*, 59(2):611–625. `https://doi.org/10.1007/s11187-021-00553-x`.

Colombo, M. G., D'Adda, D., and Pirelli, L. H. (2016). The participation of new technology-based firms in eu-funded r&d partnerships: The role of venture capital. *Research Policy*, 45(2):361–375. `https://doi.org/10.1016/j.respol.2015.10.011`.

Commission of the European Communities (2002). The european research area: providing new momentum. strengthening - reorienting - opening up new perspectives. 565 final, Brussels, 16 October.

Commission of the European Communities (2007). Green paper on the european research area: New perspectives. 161 final, Brussels, 4 April.

Crawford, G. C., Aguinis, H., Lichtenstein, B., Davidsson, P., and McKelvey, B. (2015). Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing*, 30(5):696–713. https://doi.org/10.1016/j.jbusvent.2015.01.001.

Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, complex systems*, 1695(5):1–9.

Cunningham, B. J., Gerrard, P., Schoch, H., and Lai Hong, C. (2002). An entrepreneurial logic for the new economy. *Management Decision*, 40(8):734–744. https://doi.org/10.1108/00251740210437707.

de Arroyabe, J. C. F., Schumann, M., Sena, V., and Lucas, P. (2021). Understanding the network structure of agri-food FP7 projects: An approach to the effectiveness of innovation systems. *Technological Forecasting and Social Change*, 162:120372. https://doi.org/10.1016/j.techfore.2020.120372.

De Noni, I., Orsi, L., and Belussi, F. (2018). The role of collaborative networks in supporting the innovation performances of lagging-behind european regions. *Research Policy*, 47(1):1–13. https://doi.org/10.1016/j.respol.2017.09.006.

Defazio, D., Lockett, A., and Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. *Research Policy*, 38(2):293–305. https://doi.org/10.1016/j.respol.2008.11.008.

Diestre, L. and 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.

Doherty, A. M. (2009). Market and partner selection processes in international retail franchising. *Journal of Business Research*, 62(5):528–534. https://doi.org/10.1016/j.jbusres.2008.06.011.

Ebadi, A. and Schiffauerova, A. (2015). How to receive more funding for your research? Get connected to the right people! *PloS One*, 10(7):e0133061. https://doi.org/10.1371/journal.pone.0133061.

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. and 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.

Enkel, E., Gassmann, O., and Chesbrough, H. (2009). Open R&D and open innovation: exploring the phenomenon. *R&D Management*, 39(4):311–316. https://doi.org/10.1111/j.1467-9310.2009.00570.x.

Erdil, E., Akçomak, I. S., and Çetinkaya, U. Y. (2022). Is there knowledge convergence among european regions? evidence from the european union framework programmes. *Journal of the Knowledge Economy*, 13(2):1243–1267. https://doi.org/10.1007/s13132-021-00754-5.

Erdős, P. and Rényi, A. (1959). On random graph. *Publicationes Mathematicate*, 6:290–297.

Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–60.

European Commission (2014). Horizon 2020: The New EU Framework Programme for Research and Innovation.

Everett, M. G. (2016). Centrality and the dual-projection approach for two-mode social network data. *Methodological Innovations*, 9:2059799116630662. https://doi.org/10.1177/2059799116630662.

Everett, M. G. and Borgatti, S. P. (2013). The dual-projection approach for two-mode networks. *Social Networks*, 35(2):204–210. https://doi.org/10.1016/j.socnet.2012.05.004.

Feld, B. (2012). *Startup communities: building an entrepreneurial ecosystem in your city*. John Wiley & Sons, Inc.

Feldman, M. P. (2001). The enterpreneurial event revisited: Firm formation in a regional context. *Industrial and Corporate Change*, 10(4):861–891. https://doi.org/10.1093/icc/10.4.861.

Ferraro, G. and Iovanella, A. (2016). Revealing correlations between structure and innovation attitude in inter-organisational innovation networks. *International Journal of Computational Economics and Econometrics*, 6(1):113. https://doi.org/10.1504/IJCEE.2016.073364.

Ferraro, G. and Iovanella, A. (2017). Technology transfer in innovation networks: An empirical study of the enterprise europe network. *International Journal of Engineering Business Management*, 9:1847979017735748. https://doi.org/10.1177/1847979017735748.

Foster, G., Shimizu, C., Ciesinski, S., Davila, A., Hassan, S., Jia, N., and Morris, R. (2013). Entrepreneurial Ecosystems Around the Globe and Company Growth Dynamics. In *World Economic Forum*, volume 11, pages 1–36.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239.

Friedrich, R., Peinke, J., Sahimi, M., and Tabar, M. R. R. (2011). Approaching complexity by stochastic methods: From biological systems to turbulence. *Physics Reports*, 506(5):87–162. https://doi.org/10.1016/j.physrep.2011.05.003.

Funtowicz, S. and Ravetz, J. R. (1994). Emergent Complex Systems. *Futures*, 26(6):568–582. https://doi.org/10.1016/0016-3287(94)90029-9.

Fuster, E., Padilla-Meléndez, A., Lockett, N., and Del-Águila-Obra, A. R. (2019). The emerging role of university spin-off companies in developing regional entrepreneurial university ecosystems: The case of Andalusia. *Technological Forecasting and Social Change*, 141:219–231. https://doi.org/10.1016/j.techfore.2018.10.020.

Gattringer, R., Wiener, M., and Strehl, F. (2017). The challenge of partner selection in collaborative foresight projects. *Technological Forecasting and Social Change*, 120:298–310. https://doi.org/10.1016/j.techfore.2017.01.018.

Giuliani, E. (2013). Network dynamics in regional clusters: Evidence from chile. *Research Policy*, 42(8):1406–1419. https://doi.org/10.1016/j.respol.2013.04.002.

Gloor, P. A., Dorsaz, P., Fuehres, H., and Vogel, M. (2013). Choosing the right friends - predicting success of startup entrepreneurs and innovators through their online social network structure. *International Journal of Organisational Design and Engineering*, 3(1):67–85. https://doi.org/10.1504/ijode.2013.053668.

Granstrand, O. and Holgersson, M. (2020). Innovation ecosystems: A conceptual review and a new definition. *Technovation*, 90-91:102098. https://doi.org/10.1016/j.technovation.2019.102098.

Grassi, R., Calderoni, F., Bianchi, M., and Torriero, A. (2019). Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach. *Social Networks*, 56:23–32. https://doi.org/10.1016/j.socnet.2018.08.001.

Greve, A. and Salaff, J. W. (2003). Social Networks and Entrepreneurship. *Entrepreneurship Theory and Practice*, 28(1):1–22. https://doi.org/10.1111/1540-8520.00029.

Gudivada, V. N., Rao, D., and Raghavan, V. V. (2015). Big data driven natural language processing research and applications. In *Handbook of statistics*, volume 33, pages 203–238. Elsevier. https://doi.org/10.1016/B978-0-444-63492-4.00009-5.

Guertler, M. R. and Sick, N. (2021). Exploring the enabling effects of project management for SMEs in adopting open innovation–A framework for partner search and selection in open innovation projects. *International Journal of Project Management*, 39(2):102–114. https://doi.org/10.1016/j.ijproman.2020.06.007.

Haarhaus, T., Strunk, G., and Liening, A. (2020). Assessing the complex dynamics of entrepreneurial ecosystems: A nonstationary approach. *Journal of Business Venturing Insights*, 14:e00194. https://doi.org/10.1016/j.jbvi.2020.e00194.

Hakala, H., O'Shea, G., Farny, S., and Luoto, S. (2020). Re-storying the business, innovation and entrepreneurial ecosystem concepts: The model-narrative review method. *International Journal of Management Reviews*, 22(1):10–32. https://doi.org/10.1111/ijmr.12212.

Han, J., Ruan, Y., Wang, Y., and Zhou, H. (2021). Toward a complex adaptive system: The case of the zhongguancun entrepreneurship ecosystem. *Journal of business research*, 128:537–550. https://doi.org/10.1016/j.jbusres.2019.11.077.

Hanaki, N., Nakajima, R., and Ogura, Y. (2010). The dynamics of r&d network in the it industry. *Research Policy*, 39(3):386–399. https://doi.org/10.1016/j.respol.2010.01.001.

Hassanien, A. E., Azar, A. T., Snasael, V., Kacprzyk, J., and Abawajy, J. H. (2015). Big data in complex systems. In *SBD*, volume 9. Springer. https://doi.org/10.1007/978-3-319-11056-1.

Heimann, M. and Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, 451(7176):289–292. https://doi.org/10.1038/nature06591.

Heller-Schuh, B., Barber, M., Henriques, L. M., Paier, M., Pontikakis, D., Scherngell, T., Veltri, G., Weber, M., et al. (2011). Analysis of Networks in European Framework Programmes (1984-2006). Technical report, Joint Research Centre, Seville site.

Heringa, P. W., Hessels, L. K., and 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.

Hinings, B., Gegenhuber, T., and Greenwood, R. (2018). Digital innovation and transformation: An institutional perspective. *Information and Organization*, 28(1):52–61. https://doi.org/10.1016/J.INFOANDORG.2018.02.004.

Hoang, H. and Rothaermel, F. T. (2005). The Effect of General and Partner-Specific Alliance Experience on Joint R&D Project Performance. *Academy of Management Journal*, 48(2):332–345. `https://doi.org/10.5465/AMJ.2005.16928417`.

Hoang, H. and Rothaermel, F. T. (2010). Leveraging internal and external experience: exploration, exploitation, and R&D project performance. *Strategic Management Journal*, 31(7):734–758. `https://doi.org/10.1002/smj.834`.

Hoekman, J., Scherngell, T., Frenken, K., and Tijssen, R. (2013). Acquisition of European research funds and its effect on international scientific collaboration. *Journal of Economic Geography*, 13(1):23–52. `https://doi.org/10.1093/jeg/lbs011`.

Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press.

Hu, W., Gong, Z., U, L. H., and Guo, J. (2015). Identifying influential user communities on the social network. *Enterprise Information Systems*, 9(7):709–724. `https://doi.org/10.1080/17517575.2013.804586`.

Huizingh, E. K. (2011). Open innovation: State of the art and future perspectives. *Technovation*, 31(1):2–9. `https://doi.org/10.1016/j.technovation.2010.10.002`.

Humphries, M. D. and Gurney, K. (2008). Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence. *PLoS ONE*, 3(4):e0002051. `https://doi.org/10.1371/journal.pone.0002051`.

Isenberg, D. J. (2011). The entrepreneurship ecosystem strategy as a new paradigm for economic policy: Principles for cultivating entrepreneurship. *Presentation at the Institute of International and European Affairs*, 1(781):1–13.

Isenberg, D. J. (2016). Applying the Ecosystem Metaphor to Entrepreneurship: Uses and Abuses. *The Antitrust Bulletin*, 61(4):564–573. `https://doi.org/10.1177/0003603X16676162`.

Ives, A. R. and Carpenter, S. R. (2007). Stability and diversity of ecosystems. *Science*, 317(5834):58–62. `https://doi.org/10.1126/science.1133258`.

Jackson, D. J. (2011). What is an Innovation Ecosystem? *National Science Foundation*, 1(2):1–13.

Jacobides, M. G., Cennamo, C., and Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, 39(8):2255–2276. `https://doi.org/10.1002/smj.2904`.

Jakobsen, S., Lauvås, T. A., and Steinmo, M. (2019). Collaborative dynamics in environmental r&d alliances. *Journal of Cleaner Production*, 212:950–959. https://doi.org/10.1016/j.jclepro.2018.11.285.

Jiang, R. J., Tao, Q. T., and Santoro, M. D. (2010). Alliance portfolio diversity and firm performance. *Strategic management journal*, 31(10):1136–1144. https://doi.org/10.1002/smj.869.

Kang, R. and Zaheer, A. (2018). Determinants of alliance partner choice: Network distance, managerial incentives, and board monitoring. *Strategic Management Journal*, 39(10):2745–2769. https://doi.org/10.1002/smj.2930.

Kearney, M. W. (2019). Rtweet: Collecting and analyzing twitter data. *Journal of Open Source Software*, 4(42):1829. https://doi.org/10.21105/joss.01829.

König, M. D., Liu, X., and Zenou, Y. (2019). R&D networks: Theory, empirics, and policy implications. *Review of Economics and Statistics*, 101(3):476–491. https://doi.org/10.1162/rest_a_00762.

Kumar, P. and Zaheer, A. (2019). Ego-network stability and innovation in alliances. *Academy of Management Journal*, 62(3):691–716. https://doi.org/10.5465/amj.2016.0819.

Leendertse, J., Schrijvers, M., and Stam, E. (2021). Measure Twice, Cut Once: Entrepreneurial Ecosystem Metrics. *Research Policy*, page 104336. https://doi.org/10.1016/J.RESPOL.2021.104336.

Lepori, B., Veglio, V., Heller-Schuh, B., Scherngell, T., and Barber, M. (2015). Participations to European Framework Programs of higher education institutions and their association with organizational characteristics. *Scientometrics*, 105(3):2149–2178. https://doi.org/10.1007/s11192-015-1768-2.

Lewis, T. G. (2009). Network science. *New Jersey: Wiley and Sons*, 10:9780470400791.

Li, D., Eden, L., Hitt, M. A., and Ireland, R. D. (2008). Friends, acquaintances, or strangers? Partner selection in R&D alliances. *Academy of Management Journal*, 51(2):315–334. https://doi.org/10.5465/amj.2008.31767271.

Li, F., Liu, W., and Bi, K. (2021). Exploring and visualizing spatial-temporal evolution of patent collaboration networks: A case of china's intelligent manufacturing equipment industry. *Technology in Society*, 64:101483. https://doi.org/10.1016/j.techsoc.2020.101483.

Liu, X., Li, D., Ma, M., Szymanski, B. K., Stanley, H. E., and Gao, J. (2022). Network resilience. *Physics Reports*, 971:1–108. https://doi.org/10.1016/J.PHYSREP.2022.04.002.

Ma, A., Mondragón, R. J., and Latora, V. (2015). Anatomy of funded research in science. *Proceedings of the National Academy of Sciences*, 112(48):14760–14765. https://doi.org/10.1073/pnas.1513651112.

Malecki, E. J. (2018). Entrepreneurs, networks, and economic development: A review of recent research. *Reflections and extensions on key papers of the first twenty-five years of advances (Advances in Entrepreneurship, Firm Emergence and Growth)*, 20:71–116. https://doi.org/10.1108/S1074-754020180000020010.

Mazzoni, L., Lazzeretti, L., and Innocenti, N. (2021). Entrepreneurship, complexity and the emergent order in the techno-economic scenario of the twenty-first century. evidence from a field study in tuscany. *Industry and Innovation*, 28(5):570–593. https://doi.org/10.1080/13662716.2021.1886911.

Miles, M. P. and Morrison, M. (2020). An effectual leadership perspective for developing rural entrepreneurial ecosystems. *Small Business Economics*, 54(4):933–949. https://doi.org/10.1007/S11187-018-0128-Z.

Milgram, S. (1967). The small world problem. *Psychology today*, 2(1):60–67.

Mindruta, D., Moeen, M., and Agarwal, R. (2016). A two-sided matching approach for partner selection and assessing complementarities in partners' attributes in inter-firm alliances. *Strategic Management Journal*, 37(1):206–231. https://doi.org/10.1002/smj.2448.

Minniti, M. (2008). The Role of Government Policy on Entrepreneurial Activity: Productive, Unproductive, or Destructive? *Entrepreneurship Theory and Practice*, 32(5):779–790. https://doi.org/10.1111/j.1540-6520.2008.00255.x.

Moeller, K. (2010). Partner selection, partner behavior, and business network performance: An empirical study on German business networks. *Journal of Accounting & Organizational Change*, 6(1):27–51. https://doi.org/10.1108/18325911011025687.

Moore, J. F. (1993). Predators and prey: a new ecology of competition. *Harvard Business Review*, 71(3):75–86. https://doi.org/10.1111/ijmr.12212.

Muller, E. and Peres, R. (2019). The effect of social networks structure on innovation performance: A review and directions for research. *International Journal of Research in Marketing*, 36(1):3–19. https://doi.org/10.1016/j.ijresmar.2018.05.003.

Najafian, M. and Colabi, A. M. (2014). Inter-organizational relationship and innovation: A review of literature. *Global business and management research*, 6(1):52.

Nakajima, K., Shudo, K., and Masuda, N. (2023). Higher-order rich-club phenomenon in collaborative research grant networks. *Scientometrics*, pages 1–18. https://doi.org/10.1007/s11192-022-04621-1.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied Linear Statistical Models*. Irwin Chicago.

Neumeyer, X., Santos, S. C., and Morris, M. H. (2019). Who is left out: exploring social boundaries in entrepreneurial ecosystems. *The Journal of Technology Transfer*, 44:462–484. https://doi.org/10.1007/s10961-018-9694-0.

Newman, M. (2001). Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2):025102. https://doi.org/10.1103/PhysRevE.64.025102.

Newman, M. (2002). Assortative mixing in networks. *Physical Review Letters*, 89:208701. https://doi.org/10.1103/physrevlett.89.208701.

Newman, M. (2018). *Networks*. Oxford University Press.

Newman, M. E. and Watts, D. J. (1999a). Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4-6):341–346. https://doi.org/10.1016/S0375-9601(99)00757-4.

Newman, M. E. and Watts, D. J. (1999b). Scaling and percolation in the small-world network model. *Physical review E*, 60(6):7332. https://doi.org/10.1103/PhysRevE.60.7332.

Nicotra, M., Romano, M., Del Giudice, M., and Schillaci, C. E. (2018). The causal relation between entrepreneurial ecosystem and productive entrepreneurship: A measurement framework. *The Journal of Technology Transfer*, 43:640–673. https://doi.org/10.1007/s10961-017-9628-2.

Nielsen, B. B. and Nielsen, S. (2013). Top management team nationality diversity and firm performance: A multilevel study. *Strategic Management Journal*, 34(3):373–382. https://doi.org/10.1002/smj.2021.

Nieto, M. J. and Santamaría, L. (2007). The importance of diverse collaborative networks for the novelty of product innovation. *Technovation*, 27(6-7):367–377. https://doi.org/10.1016/J.TECHNOVATION.2006.10.001.

Olanrewaju, A.-S. T., Hossain, M. A., Whiteside, N., and Mercieca, P. (2020). Social media and entrepreneurship research: A literature review. *International Journal of Information Management*, 50:90–110. https://doi.org/10.1016/J.IJINFOMGT.2019.05.011.

Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.

Paier, M. and 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.

Palla, G., Barabási, A.-L., and Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446(7136):664–667. https://doi.org/10.1038/nature05670.

Panetti, E., Parmentola, A., Ferretti, M., and Reynolds, E. B. (2020). Exploring the relational dimension in a smart innovation ecosystem: a comprehensive framework to define the network structure and the network portfolio. *The Journal of Technology Transfer*, 45(6):1775–1796. https://doi.org/10.1007/S10961-019-09735-Y.

Park, H. W. and Leydesdorff, L. (2010). Longitudinal trends in networks of university–industry–government relations in south korea: The role of programmatic incentives. *Research Policy*, 39(5):640–649. https://doi.org/10.1016/j.respol.2010.02.009.

Peel, L., Larremore, D. B., and Clauset, A. (2017). The ground truth about metadata and community detection in networks. *Science Advances*, 3(5):e1602548. https://doi.org/10.1126/sciadv.1602548.

Perotti, V. and Yu, Y. (2015). Startup Tribes: Social Network Ties that Support Success in New Firms.

Pustovrh, A., Rangus, K., and Drnovšek, M. (2020). The role of open innovation in developing an entrepreneurial support ecosystem. *Technological Forecasting and Social Change*, 152:119892. https://doi.org/10.1016/j.techfore.2019.119892.

R Core Team (2014). R: A language and environment for statistical computing.

Renner, C., Peinke, J., and Friedrich, R. (2001). Experimental indications for markov properties of small-scale turbulence. *Journal of Fluid Mechanics*, 433:383–409. https://doi.org/10.1017/S0022112001003597.

Reuer, J. J. and 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.

Roediger-Schluga, T. and Barber, M. J. (2008). R&D collaboration networks in the European Framework Programmes: Data processing, network construction and selected results. *International Journal of Foresight and Innovation Policy*, 4(3-4):321–347. https://doi.org/10.1504/IJFIP.2008.017583.

Roundy, P. T. (2017). Hybrid organizations and the logics of entrepreneurial ecosystems. *International Entrepreneurship and Management Journal*, 13(4):1221–1237. https://doi.org/10.1007/s11365-017-0452-9.

Roundy, P. T., Bradshaw, M., and Brockman, B. K. (2018). The emergence of entrepreneurial ecosystems: A complex adaptive systems approach. *Journal of Business Research*, 86:1–10. https://doi.org/10.1016/j.jbusres.2018.01.032.

Russell, M. G. and Smorodinskaya, N. V. (2018). Leveraging complexity for ecosystemic innovation. *Technological Forecasting & Social Change*, 136:114–131. https://doi.org/10.1016/j.techfore.2017.11.024.

Ryu, W., Reuer, J. J., and Brush, T. H. (2020). The effects of multimarket contact on partner selection for technology cooperation. *Strategic Management Journal*, 41(2):267–289. https://doi.org/10.1002/smj.3106.

Sarkar, M., Aulakh, P. S., and Madhok, A. (2009). Process capabilities and value generation in alliance portfolios. *Organization Science*, 20(3):583–600. https://doi.org/10.1287/orsc.1080.0390.

Scherngell, T. and 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.

Scherngell, T. and Lata, R. (2013). Towards an integrated European Research Area? Findings from Eigenvector spatially filtered spatial interaction models using European Framework Programme data. *Papers in Regional Science*, 92(3):555–577. https://doi.org/10.1111/j.1435-5957.2012.00419.x.

Schilling, M. A. and 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.

Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., and White, D. R.

(2009). Economic networks: The new challenges. *science*, 325(5939):422–425. https://doi.org/10.1126/science.1173644.

Scott, J. and Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. SAGE publications.

Scott, S., Hughes, M., and Ribeiro-Soriano, D. (2021). Towards a network-based view of effective entrepreneurial ecosystems. *Review of Managerial Science*, pages 1–31. https://doi.org/10.1007/s11846-021-00440-5.

Seetharaman, P. (2020). Business models shifts: Impact of covid-19. *International Journal of Information Management*, 54:102173. https://doi.org/10.1016/j.ijinfomgt.2020.102173.

Seo, R. (2020). Interorganizational learning for R&D consortium performance: a social capital perspective. *Journal of Knowledge Management*, 24(2):395–414. https://doi.org/10.1108/JKM-06-2019-0265.

Seppä, M. and Tanev, S. (2011). The Future of Co-Creation. *Open Source Business Resource*, Retrieved from http://www.osbr.ca/ojs/index.php/osbr/article/view/1287:6–12.

Shah, R. H. and Swaminathan, V. (2008). Factors influencing partner selection in strategic alliances: The moderating role of alliance context. *Strategic Management journal*, 29(5):471–494. https://doi.org/10.1002/smj.656.

Snijders, T. A. (2017). Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and Its Application*, 4:343–363. https://doi.org/10.1146/annurev-statistics-060116-054035.

Spender, J.-C., Corvello, V., Grimaldi, M., and Rippa, P. (2017). Startups and open innovation: a review of the literature. *European Journal of Innovation Management*, 20(1):4–30. https://doi.org/10.1108/EJIM-12-2015-0131.

Spigel, B. (2017). The Relational Organization of Entrepreneurial Ecosystems. *Entrepreneurship Theory and Practice*, 41(1):49–72. https://doi.org/10.1111/etap.12167.

Spigel, B. (2022). Examining the cohesiveness and nestedness entrepreneurial ecosystems: evidence from British FinTechs. *Small Business Economics*, pages 1–19. https://doi.org/10.1007/s11187-021-00589-z.

Spigel, B. and Harrison, R. (2018). Toward a process theory of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, 12(1):151–168. https://doi.org/10.1002/sej.1268.

Stam, E. (2015). Entrepreneurial Ecosystems and Regional Policy: A Sympathetic Critique. *European Planning Studies*, 23(9):1759–1769. https://doi.org/10.1080/09654313.2015.1061484.

Stam, E. and Van de Ven, A. H. (2021). Entrepreneurial ecosystem elements. *Small Business Economics*, 56(2):809–832. https://doi.org/10.1007/s11187-019-00270-6.

Stangler, D. and Bell-Masterson, J. (2015). Measuring an entrepreneurial ecosystem. *Available at SSRN 2580336*. https://doi.org/10.2139/SSRN.2580336.

Stanley, J. and Briscoe, G. (2010). The ABC of Digital Business Ecosystems. *arXiv preprint arXiv:1005.1899*. https://doi.org/10.48550/arxiv.1005.1899.

Szücs, F. (2020). Do research subsidies crowd out private r&d of large firms? evidence from european framework programmes. *Research Policy*, 49(3):103923. https://doi.org/10.1016/j.respol.2020.103923.

Tajudeen, F. P., Jaafar, N. I., and Ainin, S. (2018). Understanding the impact of social media usage among organizations. *Information and Management*, 55(3):308–321. https://doi.org/10.1016/j.im.2017.08.004.

Tan, F. T. C., Ondrus, J., Tan, B., and Oh, J. (2020). Digital transformation of business ecosystems: Evidence from the Korean pop industry. *Information Systems Journal*, 30(5):866–898. https://doi.org/10.1111/isj.12285.

Tatarynowicz, A., Sytch, M., and Gulati, R. (2016). Environmental demands and the emergence of social structure: Technological dynamism and interorganizational network forms. *Administrative Science Quarterly*, 61(1):52–86. https://doi.org/10.1177/0001839215609083.

Thurner, S., Hanel, R., and Klimek, P. (2018). *Introduction to the theory of complex systems*. Oxford University Press.

Tsai, K.-H. (2009). Collaborative networks and product innovation performance: Toward a contingency perspective. *Research Policy*, 38(5):765–778. https://doi.org/10.1016/j.respol.2008.12.012.

Tsujimoto, M., Kajikawa, Y., Tomita, J., and Matsumoto, Y. (2018). A review of the ecosystem concept — Towards coherent ecosystem design. *Technological Forecasting & Social Change*, 136:49–58. https://doi.org/10.1016/j.techfore.2017.06.032.

Tumasjan, A., Braun, R., and Stolz, B. (2021). Twitter sentiment as a weak signal in venture capital financing. *Journal of Business Venturing*, 36(2):106062. https://doi.org/10.1016/J.JBUSVENT.2020.106062.

Valente, T. W., Coronges, K., Lakon, C., and Costenbader, E. (2008). How correlated are network centrality measures? *Connections (Toronto, Ont.)*, 28(1):16.

Van de Ven, A. H. (1986). Central problems in the management of innovation. *Management Science*, 32(5):590–607. https://doi.org/10.2307/2631848.

Van de Ven, A. H. (1993). The development of an infrastructure for entrepreneurship. *Journal of Business Venturing*, 8(3):211–230. https://doi.org/10.1016/0883-9026(93)90028-4.

Vătămănescu, E.-M., Cegarra-Navarro, J.-G., Andrei, A. G., Dincă, V.-M., and Alexandru, V.-A. (2020). SMEs strategic networks and innovative performance: a relational design and methodology for knowledge sharing. *Journal of Knowledge Management*, 24(6):1369–1392. https://doi.org/10.1108/JKM-01-2020-0010.

Wagner, C. S. and Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, 34(10):1608–1618. https://doi.org/10.1016/j.respol.2005.08.002.

Wang, Y. and Rajagopalan, N. (2015). Alliance capabilities: Review and research agenda. *Journal of Management*, 41(1):236–260. https://doi.org/10.1177/0149206314557157.

Wanzenböck, I., Lata, R., and 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.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*, volume 8. Cambridge University Press.

Wathne, K. H., Heide, J. B., Mooi, E. A., and Kumar, A. (2018). Relationship governance dynamics: The roles of partner selection efforts and mutual investments. *Journal of Marketing Research*, 55(5):704–721. https://doi.org/10.1177/0022243718801325.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442. https://doi.org/10.1038/30918.

Weng, C. S., Yang, W.-G., and Lai, K.-K. (2014). Technological position in alliances network. *Technology Analysis & Strategic Management*, 26(6):669–685. https://doi.org/10.1080/09537325.2014.923096.

West, D. B. et al. (2001). *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River.

Whetsell, T. A. (2023). Democratic governance and global science: A longitudinal analysis of the international research collaboration network. *Plos One*, 18(6):e0287058. `https://doi.org/10.1371/journal.pone.0287058`.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 48(4):817–838. `https://doi.org/10.2307/1912934`.

Williams, O. E., Mazzarisi, P., Lillo, F., and Latora, V. (2022). Non-markovian temporal networks with auto-and cross-correlated link dynamics. *Physical Review E*, 105(3):034301. `https://doi.org/10.1103/PhysRevE.105.034301`.

Wu, G., Zhao, X., Zuo, J., and Zillante, G. (2019). Effects of team diversity on project performance in construction projects. *Engineering, Construction and Architectural Management*, 26(3):408–423. `https://doi.org/10.1108/ECAM-05-2018-0220`.

Wu, W. Y., Shih, H.-A., and Chan, H.-C. (2009). The analytic network process for partner selection criteria in strategic alliances. *Expert Systems with Applications*, 36(3):4646–4653. `https://doi.org/10.1016/j.eswa.2008.06.049`.

Wurth, B., Stam, E., and Spigel, B. (2022). Toward an entrepreneurial ecosystem research program. *Entrepreneurship Theory and Practice*, 46(3):729–778. `https://doi.org/10.1177/1042258721998948`.

Xie, Q. and Su, J. (2021). The spatial-temporal complexity and dynamics of research collaboration: Evidence from 297 cities in china (1985–2016). *Technological Forecasting and Social Change*, 162:120390. `https://doi.org/10.1016/j.techfore.2020.120390`.

Yue, C. A., Thelen, P., Robinson, K., and Men, L. R. (2019). How do ceos communicate on twitter? a comparative study between fortune 200 companies and top startup companies. *Corporate Communications: An International Journal*, 24(3):532–552. `https://doi.org/10.1108/CCIJ-03-2019-0031`.

Zhang, Y. and Chen, K. (2022). Network growth dynamics: The simultaneous interaction between network positions and research performance of collaborative organisations. *Technovation*, 115:102538. `https://doi.org/10.1016/j.technovation.2022.102538`.

Zhang, Y., Chen, K., Zhu, G., Yam, R. C., and Guan, J. (2016). Inter-organizational scientific collaborations and policy effects: an ego-network evolutionary perspective of the chinese academy of sciences. *Scientometrics*, 108:1383–1415. `https://doi.org/10.1007/s11192-016-2022-2`.

Zirulia, L. (2023). Path dependence in evolving r&d networks. *Journal of Evolutionary Economics*, 33(1):149–177. `https://doi.org/10.1007/s00191-022-00802-6`.

Appendix 1.A for: "Network-based principles of entrepreneurial ecosystems: a case study of a start-up network"

Data collection

Step 1

The lists of the most promising start-ups for each country in 2019/2020 are collected from accredited web sources which provide every year the names of the most noteworthy start-ups in the international context. In particular, these data sources indicate fast-growing start-ups according to various criteria, including funding raised, revenues generated, capital invested, number of employees, industry, and business model. All the sources are listed below (countries for which the source is used are specified in parentheses):

- *Sifted*[2][3][4][5][6] (Austria, France, Germany, Italy): *Sifted* is an online publication backed by *The Financial Times* that provides in-depth reporting on the European start-up environment, including company updates, funding rounds, venture capital deals, founder stories and "how-to's".

- *Seedtable*[7][8][9] (Austria, France, Germany): *Seedtable* provides an analysis of the European tech sector and its impact on society. Its audience comprises more than 10,000 founders, employees and investors.

[2] https://sifted.eu/italy-startups-top-rankings/ (Date accessed: 01/04/2020)
[3] https://sifted.eu/german-startups-top-rankings/ (Date accessed: 22/03/2020)
[4] https://sifted.eu/german-startups-top-rankings/ (Date accessed: 22/03/2020)
[5] https://sifted.eu/french-startups-top-rankings/ (Date accessed: 23/03/2020)
[6] https://sifted.eu/austria-startups-top-rankings/ (Date accessed: 30/03/2020)
[7] https://www.seedtable.com/startups-germany (Date accessed: 20/03/2020)
[8] https://www.seedtable.com/startups-france (Date accessed: 23/03/2020)
[9] https://www.seedtable.com/startups-austria (Date accessed: 31/03/2020)

- *StartupItalia*[10] (Italy): *StartupItalia* is an Italian innovation magazine that tells entrepreneurial stories, highlighting both difficulties and successes of novel start-ups, providing them precious advices. *StartupItalia* realizes a list of the most interesting Italian start-ups every year in collaboration with *Bocconi University*.

- *Top 100 Swiss Startup Award*[11] (Switzerland): The *Top 100 Swiss Startup Award* is the yearly reference event where the most promising Swiss startups' CEOs, investors, and corporate executives gain deep insights into the latest proven concepts and establish new business relationships. Every year, the 100 most innovative and promising Swiss start-ups are selected by a jury of 100 leading investors and start-up experts.

- *Startups*[12] (United Kingdom): *Startups* is a key resource for starting a business in the UK, helping start-ups from the origin of business ideas to exiting out. The *Startups 100* has become an annual milestone in the UK business calendar, highlighting up-and-coming businesses with the potential to change their industries for the better.

- *StartupsReal by El Referente*[13] (Spain): *StartupsReal by El Referente* is Spain's first digital daily magazine specialized in start-ups and innovation. The media group has also recently launched *StartupsReal*, the first English-language magazine dedicated to start-ups and entrepreneurs in Spain. They indicate the start-ups with the most potential to become Spain's future scale-ups.

Step 2

All the start-ups identified during the previous step are classified in fifteen different categories based on their business activities, according to the criteria defined in Table 2.

[10] https://startupitalia.eu/119441-20191213-le-top-100-le-migliori-startup-italiane-del-2019-secondo-noi (Date accessed: 12/03/2020)
[11] https://www.top100startups.swiss/index.cfm?page1̄36340&cfid3̄50485093&cftoken7̄0fdd71bcc1f822d-2D69F011-CA87-A39D-E9BC482C73E7E3E0& (Date accessed: 15/03/2020)
[12] https://startups.co.uk/startups-100/2019/startups-100-2019/ (Date accessed: 17/03/2020)
[13] https://www2.deloitte.com/content/dam/Deloitte/es/Documents/acerca-de-deloitte/Deloitte-ES-aboutdeloitte-ranking100-startups.pdf (Date accessed: 26/03/2020)

Table 2: Startups' categories.

Category	Definition
Agritech	Startups aiming to reshape global agriculture by increasing the productivity of the agriculture system while reducing the environmental and social costs associated with current production practices.
Biotech	Startups operating in the broad area of biology, involving living systems and organisms, to develop or make products. Biotechnology encompasses emerging sciences, including genomics, recombinant gene techniques, applied immunology, and the advancement of pharmaceutical therapies and diagnostic tests.
Cleantech	Startups engaged in any process or producing any product that mitigate negative environmental impacts through substantial improvements in energy efficiency, sustainable resource utilization, or environmental protection activities.
Electronics, mechanics	Startups involved in the production of innovative electronic devices, including those realted to augmented and virtual reality, or startups specialized in mechanical equipment.
Fintech	Startups offering innovative banking services or engaging in cryptocurrency operations.
Insurtech	Startups providing innovative insurance services and policies across various fields.
Internet, mobile	Startups with a business model centered around marketplaces; startups providing media platforms; startups offering mobile services. This category is broader than the others due to a wider identification criterion.
Legaltech	Startups offering legal expertise through experienced lawyers and judges.
Materials, chemicals	Startups developing new types of materials or working on chemical products.
Medtech	Startups leveraging technology in healthcare settings, including disposables, capital equipment, and surgical procedures, through implant technology, biomaterials, and connected health IT.
Micro/nano technology	Startups developing micro and nano tachnologies.
Mobility	Startups operating in the mobility sector and providing services of vehicle-sharing.
Proptech	Startups using information technology to assist individuals and companies in searching, buying, selling, and managing real estate.
Software	Startups developing enterprise softwares and management platforms.
Others	Startups that do not belong to any of the previous categories.

Appendix 2.A for: "How do partner selection strategies affect the amount of funds in collaborative research projects? Evidence using the dual-projection approach"

Proof of Proposition 1

By a well-known result of linear algebra, $\mathbf{A}$ can be expressed using its spectral decomposition:

$$\mathbf{A} = \sum_{j=1}^{n} \lambda_j \mathbf{x}_j \mathbf{x}_j^T$$

In general, it holds:

$$\mathbf{A}^k = \sum_{j=1}^{n} \lambda_j^k \mathbf{x}_j \mathbf{x}_j^T \quad \forall k \in \mathbb{N} \tag{13}$$

By Eq. (4) and Eq. (7) we have the following equalities:

$$\mathbf{c}\left(\beta\right) = \frac{1}{\beta} \sum_{k=1}^{\infty} \left(\beta^k \sum_{j=1}^{n} \lambda_j^k \mathbf{x}_j \mathbf{x}_j^T \right) \mathbf{1} = \frac{1}{\beta} \sum_{j=1}^{n} \left(\sum_{k=1}^{\infty} (\beta\lambda_j)^k \right) \mathbf{x}_j \mathbf{x}_j^T \mathbf{1}$$

If $|\beta\lambda_j| < 1 \quad \forall j = 1, ..., n$, then the series is convergent. We thus assume the more restrictive condition $0 < \beta < \frac{1}{\lambda_1}$, that ensures the convergence to $\frac{\beta\lambda_j}{1-\beta\lambda_j} \quad \forall j = 1, ..., n$. Then:

$$\mathbf{c}(\beta) = \sum_{j=1}^{n} \frac{\lambda_j}{1 - \beta\lambda_j} \mathbf{x}_j \mathbf{x}_j^T \mathbf{1}.$$

Notice that, when β approaches 0^+, $\mathbf{c}\,(\beta) = \sum_{j=1}^{n} \lambda_j \mathbf{x}_j \mathbf{x}_j^T \mathbf{1}$ $=\mathbf{A1} = \mathbf{d}$. Moreover:

$$\mathbf{c}(\beta) = \frac{\lambda_1}{1-\beta\lambda_1}\mathbf{x}_1\mathbf{x}_1^T\mathbf{1} + \sum_{j=2}^{n} \frac{\lambda_j}{1-\beta\lambda_j}\mathbf{x}_j\mathbf{x}_j^T\mathbf{1}.$$

When β approaches $\frac{1}{\lambda_1}^{-}$, the first term $\frac{\lambda_1}{1-\beta\lambda_1}\mathbf{x}_1\mathbf{x}_1^T\mathbf{1}$ prevails on the others in the sum, then $\mathbf{c}(\beta) \approx \left(\frac{\lambda_1}{1-\beta\lambda_1}\right)\mathbf{x}_1\mathbf{x}_1^T\mathbf{1} = \left(\frac{\lambda_1\mathbf{x}_1^T\mathbf{1}}{1-\beta\lambda_1}\right)\mathbf{x}_1$. Therefore, the Beta-centrality $\mathbf{c}(\beta)$ can be considered proportional to the eigenvector centrality $\mathbf{x}_1$.

Proof of Proposition 2

At first, we recall that $\mathbf{c}(\beta)^{(p)}_{dual} = \mathbf{E}^T\mathbf{c}(\beta)^{(o)}$, where $\mathbf{c}(\beta)^{(o)}$ is the vector of Beta-centrality of matrix $\mathbf{EE}^T$. By Eq. (2.4), assuming $0 < \beta < \frac{1}{\lambda_1}$, and expressing $\mathbf{EE}^T$ by its spectral decomposition, i.e., $\mathbf{EE}^T = \sum_{j=1}^{n_2} \lambda_j \mathbf{x}_j \mathbf{x}_j^T$, we have:

$$\begin{aligned}
\mathbf{c}(\beta)^{(p)}_{dual} = \mathbf{E}^T\mathbf{c}(\beta)^{(o)} = \mathbf{E}^T \sum_{k=1}^{\infty} \beta^{k-1}(\mathbf{EE}^T)^k\mathbf{1} = \\
\frac{1}{\beta}\mathbf{E}^T \sum_{k=1}^{\infty} \beta^{k}(\mathbf{EE}^T)^k\mathbf{1} = \frac{1}{\beta}\mathbf{E}^T \sum_{k=1}^{\infty} \left(\beta^k \sum_{j=1}^{n_2} \lambda_j^k \mathbf{x}_j\mathbf{x}_j^T\right)\mathbf{1} = \\
\frac{1}{\beta}\mathbf{E}^T \sum_{j=1}^{n_2}\sum_{k=1}^{\infty} (\beta\lambda_j)^k \mathbf{x}_j\mathbf{x}_j^T\mathbf{1} = \frac{1}{\beta}\mathbf{E}^T \sum_{j=1}^{n_2} \frac{\beta\lambda_j}{1-\beta\lambda_j}\mathbf{x}_j\mathbf{x}_j^T\mathbf{1} = \\
\mathbf{E}^T \sum_{j=1}^{n_2} \frac{\lambda_j}{1-\beta\lambda_j}\mathbf{x}_j\mathbf{x}_j^T\mathbf{1}
\end{aligned}$$

where the convergence to $\frac{\beta\lambda_j}{1-\beta\lambda_j}$ $\forall j = 1, ..., n$ is ensured by the condition $0 < \beta < \frac{1}{\lambda_1}$. Therefore, when β approaches 0+, $\mathbf{c}(\beta)^{(p)}_{dual} = \mathbf{E}^T \sum_{j=1}^{n_2} \lambda_j \mathbf{x}_j\mathbf{x}_j^T\mathbf{1} = \mathbf{E}^T(\mathbf{EE}^T)\mathbf{1} = \mathbf{E}^T\mathbf{d}^{(o)}$, that returns, by Eq. (2.6), the dual degree centrality of projects $\mathbf{d}^{(p)}_{dual}$.

Moreover:

$$\mathbf{c}(\beta)^{(p)}_{dual} = \mathbf{E}^T\frac{\lambda_1}{1-\beta\lambda_1}\mathbf{x}_1\mathbf{x}_1^T\mathbf{1} + \mathbf{E}^T\sum_{j=2}^{n_2} \frac{\lambda_j}{1-\beta\lambda_j}\mathbf{x}_j\mathbf{x}_j^T\mathbf{1}$$

When β approaches $\frac{1}{\lambda_1}^{-}$, the first term $\mathbf{E}^T\frac{\lambda_1}{1-\beta\lambda_1}\mathbf{x}_1\mathbf{x}_1^T\mathbf{1}$ prevails on the others in the sum, then $\mathbf{c}(\beta)^{(p)}_{dual} \approx \mathbf{E}^T\left(\frac{\lambda_1}{1-\beta\lambda_1}\right)\mathbf{x}_1\mathbf{x}_1^T\mathbf{1} = \left(\frac{\lambda_1\mathbf{x}_1^T\mathbf{1}}{1-\beta\lambda_1}\right)\mathbf{E}^T\mathbf{x}_1$. Notice that $\mathbf{x}_1$ represents the eigenvector centrality of organizations, denoted by $\mathbf{x_1}^{(o)}$ by Eq. (6). Therefore, we can identify in $\mathbf{E}^T\mathbf{x}_1$ the dual eigenvector centrality of projects $\mathbf{x}^{(p)}_{1,dual}$, so that, the dual Beta-centrality $\mathbf{c}(\beta)^{(p)}_{dual}$ can be considered proportional to the dual eigenvector centrality $\mathbf{x}^{(p)}_{1,dual}$.