Determining the dynamics of collaboration in EU Framework Programmes under a network perspective

Andrea Ancona
Sapienza University of Rome, Italy
Roy Cerqueti
Sapienza University of Rome, Italy
Gianluca Vagnani
Sapienza University of Rome, Italy

ISSN 2466-7498 and ISBN 978-2-9602195-5-5

Determining the dynamics of collaboration in

EU Framework Programmes under a network perspective

Abstract

Collaborative networks gained attention in the field of economics of innovation in the recent past. One of the main interests concerns the temporal analysis of such networks, both in a scientific and in a European policy context. At the European level indeed, the objective is to promote strong and durable partnerships among research institutions and with industry, going beyond the usual project-based cooperation. The purpose of this study is to investigate these long-lasting collaborative relationships between the organizations that received funds by all the first eight European Framework Programmes (EU FPs). 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. Considering participations in European projects funded by all the first eight EU FPs gives us the possibility to analyze the dynamics of collaborations in the context of European research projects over a long-time span. In more detail, we adopt a novel approach to model the dynamics of participation in EU FPs by means of Social Network Analysis (SNA) and statistics tools. The main objective is to estimate the probabilities of moving from one position to another - in terms of centrality measures - across different FPs, and to understand if the position within subsequent collaborative research networks is affected by a certain path dependency. Our results confirm the existence of a path dependency, in the sense that participating in previous FPs provides a competitive advantage to organizations due to several network benefits, such as growing experience, competencies, and popularity. Phenomena of "preferential attachment" are also evident. Finally, we find that the estimated probability transition matrices are able to highlight

relevant events that affected the European Union and its strategies in the field of research, which are the Treaty of Maastricht and the adoption of the European Research Area (ERA).

Keywords: Collaborative networks; EU Framework Programmes; Research projects; Collaborative dynamics; Social Network Analysis.

1. Introduction

The establishment of a systematic Research and Technological Development (RTD) policy at the European level can be traced back to the 1980s in correspondence with the first EU FP. The RTD policy goes in the same direction as the broader cohesion policy, which is aimed at conceiving the EU as a common market, promoting the free circulation of people, goods, and capital, and the continuous exchange of knowledge. An essential role of the EU FPs is to provide funds for transnational networks of researchers in order to foster international research collaboration. 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 (Commission of the European Communities, 2007): an adequate flow of competent researchers; world-class research infrastructures; excellent research institutions; effective knowledge sharing; well-coordinated research programmes and priorities; a wide opening of the ERA to the world. 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. 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 this paper, we aim to investigate the dynamics of European collaborative research projects both at a macro and a micro-level. Specifically, we map the local behaviors of single actors in terms of their position in a collaborative network through the use of centrality measures. Then, we model the dynamics of collaboration among the organizations participating in projects funded by all the first eight EU FPs by means of statistical tools in order to estimate the transition probability matrices from one level of centrality to another over consecutive FPs. This approach can provide relevant insights into the evolution of European collaborative research networks, highlighting eventual phenomena of path dependency and predatory behaviors. Furthermore, this study can provide practical information to participant organizations as well about the probability of becoming core members in subsequent FPs.

In this way, we aim to enlarge the strand of literature focused on the analysis of collaborative networks, specifically in the field of European research projects. In particular, this paper represents an advancement with respect to the state of the art for different reasons. At first, unlike many other works in the literature, we do not want to aggregate organizations at the country level. Rather, the objective is to model their behavior as single entities. Secondly, to the best of our knowledge, this is the first attempt to take into account all the past EU FPs, from the first (FP1) to the eighth (Horizon 2020). This continuity along consecutive FPs allows us to model the dynamics of participation in European projects through statistical tools over a longer time horizon compared to previous studies. Moreover, previous works analyze mainly macro-dynamics and dynamics on average, while this research is based on the characterization of the micro-dynamics at the participant level. As a consequence, this work, in addition to having implications in a European policy context, aims to provide relevant indications to participant organizations as well. Indeed, networking capability and inter-organizational knowledge mechanisms have a substantial impact on firm innovation performance (Mokhtarzadeh et al., 2020). Thus, understanding how their centrality increase in a collaborative network is valuable for organizations to identify possible strategies to pursue. From a European policy perspective instead, a dynamic study of EU FPs can shed light on the functioning of the RTD policy and on the actual advancements in promoting the ERA.

The remainder of the paper is organized as follows. An overview of the state of the art on the analysis of EU FPs is provided in Section 2. We introduce the data and its source in Section 3. Section 4 describes our methodological approach, that relies on elements from SNA and statistics. Results are presented in Section 5, while Section 6 highlights the main contributions and possible further developments of this work.

2. State of the art

The analysis of projects funded by the EU FPs constitute one of the most interesting case studies in the field of collaborative R&D at the European level. Several authors laid the foundation for this research, that presents however many open issues to deal with, such as a global understanding of funding mechanisms, the criteria behind partner selection strategies, and a dynamic observation of collaborations over different FPs.

Scherngell and Barber (2011) analyze joint research projects funded by the fifth FP to determine how the variation of cross-region industry and public research is affected by geography, finding that spatial factors significantly affect industrial collaboration, while in the public sector the effect is smaller. Similarly, Scherngell and Lata (2013) use a spatial interaction model to estimate how specific separation effects influence the variation of R&D networks in Europe, analyzing projects mostly funded by the fifth and the sixth FPs. While confirming that geographical distance exerts a negative effect on collaboration probability, they show that the effect significantly decreases between 1999 and 2006. Distance, in terms of geographical, economic, technological, and social factors, matters also in the probability of two regions establishing an R&D collaboration in the seventh FP, revealing that an integrated

ERA is still far from being reached (Amoroso et al., 2018). Moreover, there appears to exist a certain path dependency between close and similar regions, with a high degree of persistence over time. Hoekman et al. (2013) investigate the link between scientific collaboration networks, represented by co-publications in scientific journals, and joint participations in the fifth and the sixth FPs between two regions, revealing that co-publishing does not have a particular impact on the amount of received funds. With data drawn from 350 project managers who are actively involved in European innovation networks, Arranz and de Arroyabe (2012) argue that the efficient design of process, structure, and governance subsystems boosts the performance of innovation networks, while the interrelations between subsystems have complementary and synergic effects. Breschi and Cusmano (2004) are among the first to examine the network of R&D joint ventures (RJVs) funded by the European Commission between 1992-1996. Their study shows the emergence of an oligarchic behavior as a consequence of previous cooperative programmes, such that a group of core actors holds the leadership of the network and allows the remaining organizations to communicate. The study of Lepori et al. (2015) analyzes the participation of higher education institutions (HEIs) in EU FPs and their association with HEI characteristics, country, and geographical effects. Their results suggest the close relation between HEI reputation and the network structure of EU FP participants. Enger and Castellacci (2016) present a timely analysis of participation in Horizon 2020, which denotes that the propensity to apply is strengthened by prior participation in other FPs and the existence of complementary national funding schemes. Furthermore, the probability of being funded is enhanced by prior participation as well as the scientific reputation of the applicant organization.

The use of SNA to address the position in European collaborative research networks is exploited by Enger (2018). He finds that HEIs with high and low levels of centrality in the seventh FP display a significantly greater propensity to apply to Horizon 2020, compared to

HEIs with no centrality. Moreover, HEIs with high levels of centrality have a significantly greater propensity to obtain funds, with respect to the group with no centrality. Even Balland et al. (2019) use centrality measures to investigate the dynamics of collaborative research networks related to the sixth, the seventh and the eighth FP. They find that participants from EU-15 countries tend to be more central than participants from EU-13 countries, as well as associated and third countries, in Horizon 2020. On average, participants are slightly less central in FP7 and Horizon 2020 compared to FP6, possibly revealing the entry of smaller players. The global structure of the networks - in terms of assortativity, inequality, degree distribution, and average path length - remains stable over the three FPs. Cinelli et al. (2022) instead, introduce a new metric, called collective network effect (CNE), to measure the benefit of network membership among the participants in projects funded by the seventh FP. They find that organizations with a higher CNE generally have access to more funds than those with a lower CNE.

3. Data

The list of projects funded by the EU FPs, and the related participant organizations, are publicly available on the Community Research and Development Information Service (CORDIS) website[1]. CORDIS is the European Commission's primary source of results about projects funded by the EU FPs, with a unique and structured repository containing information about all projects financed from FP1 to Horizon Europe (i.e., the current FP), and about the related participants. However, data from CORDIS (especially data referred to the oldest FPs) contains different sources of errors related to the use of several distinct names to address the same organization. In their work, Roediger-Schluga and Barber (2008) tackle this issue and introduce for the first time a novel data source of higher quality than the original set

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

of data, accounting for the first five EU FPs (i.e., the EUPRO database). During the following years, the EUPRO database has been expanded, including all FPs from FP1 to Horizon 2020 (H2020), as well as other European programmes such as EUREKA, Joint Technology Initiatives (JTI), and European Cooperation in Science and Technology (COST) (Heller-Schuh et al., 2020). Actually, most of the studies dealing with European projects benefit from the EUPRO database, performing their analyses on this data source source (Paier and Scherngell, 2011; Scherngell and Barber, 2011; Hoekman et al., 2013; Scherngell and Lata, 2013; Lepori et al., 2015; Crespo et al., 2016; Heringa et al., 2016; Uhlbach et al., 2017; Wanzenböck et al., 2020; Cavallaro and Lepori, 2021).

Thus, we requested and obtained the access to the EUPRO database containing the list of projects funded by the first eighth 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, then 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 1, differentiating among the eight FPs.

Framework Programme	Distinct projects	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

Table 1: EUPRO database

We decided to consider the organizations participating in all the eight FPs, in order to analyze the whole dynamics from FP1 to H2020. The final amount of selected participants is equal to 509 organizations.

4. Methodological approach

In this paper, we aim to map the local behaviors of participant organizations that received funds from FP1 to H2020 in terms of their position in the respective collaborative networks. In more detail, we analyze the evolution of their connections based on the values of their centrality measures. Then, we model the dynamics of collaborative relationships through the estimation of the transition probability matrices, whose values correspond to the probability to move from one level of centrality to another over consecutive FPs. Thus, our approach combines elements from SNA and statistics, that we introduce below.

4.1 Social Network Analysis (SNA) and collaborative networks

SNA is one of the most powerful instruments 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 structure (Hu et al, 2015) and measures of centrality (Scott and Carrington, 2011). Networks are indeed the preferred tool for mapping the interactions among the members of a system, and to describe structures, roles and dynamics of complex systems (Börner et al., 2007).

SNA has its foundation in graph theory. In fact, a network is represented by a graph $G = (V, E)$, where V is the set of vertices (also called nodes) and E is the set of edges (also called links). Two vertices i and j are adjacent if $(i, j) \in E$. The adjacency relations between nodes are described by a n-square (where n is the number of nodes of a graph) binary matrix A, that is called adjacency matrix, whose elements $a_{ij} = 1$ if $(i, j) \in E$, 0 otherwise. In case of a weighted graph, i.e., there is a weight on edges, 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} if $(i,j) \in E$, 0 otherwise.

Data about participation in collaborative projects belongs to the wide category of two-mode data. This kind of data is characterized by two distinct sets of nodes (e.g., actors and events) that are connected by a relation (e.g., "actors participate in events"), while there are no connections among nodes in the same set. In this case, the two sets of nodes are represented by organizations and projects, and they are linked through the relation "organizations participate in projects". Two-mode data is well represented in SNA by a bipartite graph. A graph is bipartite if the set V of its nodes can be partitioned into two different subsets V_1 and V_2 such that there are no links internal to $V_1$and V_2, and all edges hold between nodes in $V_1$and V_2, respectively. According to Borgatti and Halgin (2011), the most popular approach to deal with two-mode data is the "conversion" approach, i.e., two-mode data are converted into two one-mode projections, where usually just one of them is of specific interest. In our case, we work with the one-mode projected networks whose nodes correspond to the distinct participants get funded by all the first eight EU FPs, and two nodes are connected if the corresponding organizations are partners in one or more projects within the same FP. Thus, participations in projects funded by a specific FP are ultimately represented by an undirected, weighted network, where weights on edges constitute eventual multiple partnerships between the same organizations in different projects. Then, our case study develops on the analysis of eight different one-mode projected networks, each standing for a distinct FP.

The best way to assess the relevance of an organization in a collaborative network is by computing its centrality measures. Centrality is indeed one of the most relevant concepts in SNA, aimed to determine the importance of a node based on the number and the quality of its connections. There are several centrality measures that have become popular in network analysis. The most intuitive one is degree centrality, which is equal to the number of adjacent nodes of a vertex. Since our analyses are performed on weighted networks, we decide to consider as a proxy of the relevance of participant organizations, a variant of degree centrality, i.e., the strength centrality. Formally, the strength centrality (called "strength" below) of a node i is defined as:

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

Thus, the higher the amount of projects an organization is involved in, the greater its relevance within the collaborative network.

Therefore, we compute a vector of centralities for every FP, whose components correspond to the value of strength of each organization receiving funds by all the FPs from FP1 to H2020.

4.2 Transition probability matrices

Our main objective is to estimate the probability to move from one level of strength to another over consecutive FPs. For each pair of consecutive FPs, we can rely on the empirical probability matrix P, whose dimension is equal to $n \times m$, where n 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 on the rows we have the vector of distinct values of strength in FP1 $s = (s_1, \cdots, s_n)$ in ascending order, while on the columns we have the vector of distinct values of strength in FP2 $s' = (s'_1, \cdots, s'_m)$ in ascending order. Thus, P has the following structure:

$$P = \begin{pmatrix} p_{1,1} & \cdots & p_{1,m} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \cdots & p_{n,m} \end{pmatrix}$$

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, which is computed as:

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

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 values of strength in FP1 and FP2, respectively. This analysis would provide a level of detail too specific to be interpreted in a proper way. It would be much more relevant from a policy perspective to partition the values of centrality in classes (e.g., low, medium, and high) in order to estimate the probability of moving, for instance, from a low to a high level of centrality, and to understand if there exists a certain path dependency in collaborative research projects funded by the EU FPs.

To this aim, the most common approach is to determine partitions exogenously by defining either fixed or quantile thresholds, i.e., based on the distribution (e.g., the first quartile is low, the second is medium, and the last two quartiles correspond to high). Nevertheless, this approach introduces some biases in the analysis due to the authors' subjective choice of fixed or quantile thresholds. For this reason, Cerqueti et al. (2017) adopt a novel methodology to define the partition in classes for time series endogenously aimed to minimize the distance between the estimated probability to move from one level of centrality to another and the empirical transition probability from a distinct value of centrality to another. In this paper, we adapt this approach to partition the distinct values of strength in each FP, which are then grouped together according to the likelihood of obtaining partitions that evolve as similarly to the effective transitions as possible.

Specifically, we want to identify three different classes of strength for each FP, i.e., low, medium, and high. Thus, we need to determine two distinct thresholds t_1 and t_2 such that:

$$Low = \{s_i \mid s_i \leq t_1\}$$

$$Medium = \{s_i \mid t_1 < s_i \leq t_2\}$$

$$High = \{s_i \mid t_2 < s_i\}$$

Given a set of n distinct values of strength, the number of possible partitions to test is equal to $\frac{(n-1)(n-2)}{2}$. It is important to note that we partition row values for each empirical probability matrix describing the relation between two consecutive FPs. For all possible partitions, we estimate a theoretical probability matrix π, which is defined as follows:

$$\pi = \begin{pmatrix} \pi_{1,1} & \cdots & \pi_{1,m} \\ \vdots & \ddots & \vdots \\ \pi_{n,1} & \cdots & \pi_{n,m} \end{pmatrix}$$

where in correspondence of all the values of strength belonging to the same class, there is the same probability. For instance, if we partition the set of distinct values of strength in FP1 in the following way:

$$Low = \{s_1, \cdots, s_h\}$$

$$Medium = \{s_{h+1}, \cdots, s_k\}$$

$$High = \{s_{k+1}, \cdots, s_n\}$$

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

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

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

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

Therefore, we will select the best partition based on the estimated theoretical probabilities. More specifically, the final partition is the one that minimizes the distance between the empirical probability matrix P and the theoretical probability matrix π, formally:

$$\min d_{P,\pi} = \min \sum_{i=1}^{n} \sum_{j=1}^{m} |p_{i,j} - \pi_{i,j}|$$

Once the best partition is determined for FP1, the same procedure is applied to the empirical probability matrix describing the transition from FP2 and FP3, in order to define the appropriate thresholds for FP2 as well. In this way, we obtain the final transition probability matrix $\pi^{1,2}$ from FP1 to FP2, that is a 3×3 squared matrix defined as follows:

$$\pi^{1,2} = \begin{pmatrix} \pi_{1,1}^{1,2} & \pi_{1,2}^{1,2} & \pi_{1,3}^{1,2} \\ \pi_{2,1}^{1,2} & \pi_{2,2}^{1,2} & \pi_{2,3}^{1,2} \\ \pi_{3,1}^{1,2} & \pi_{3,2}^{1,2} & \pi_{3,3}^{1,2} \end{pmatrix}$$

where:

$$\pi_{1,1}^{1,2} = \sum_{j=1}^{h\prime} \pi_{1,j} \ ; \ \pi_{1,2}^{1,2} = \sum_{j=h'+1}^{k\prime} \pi_{1,j} \ ; \ \pi_{1,3}^{1,2} = \sum_{j=k'+1}^{m} \pi_{1,j}$$

$$\pi_{2,1}^{1,2} = \sum_{j=1}^{h\prime} \pi_{h+1,j} \ ; \ \pi_{2,2}^{1,2} = \sum_{j=h'+1}^{k\prime} \pi_{h+1,j} \ ; \ \pi_{2,3}^{1,2} = \sum_{j=k'+1}^{m} \pi_{h+1,j}$$

$$\pi_{3,1}^{1,2} = \sum_{j=1}^{h\prime} \pi_{k+1,j} \ ; \ \pi_{3,2}^{1,2} = \sum_{j=h'+1}^{k\prime} \pi_{k+1,j} \ ; \ \pi_{3,3}^{1,2} = \sum_{j=k'+1}^{m} \pi_{k+1,j}$$

where h' and k' are the first and the second thresholds of FP2, respectively. Here, the interpretation is that the average probability to move from a low level of strength in FP1 to a low level of strength in FP2 is $\pi_{1,1}^{1,2}$. The average probability to move from a low level of strength in FP1 to a medium level of strength in FP2 is $\pi_{1,2}^{1,2}$. The average probability to move from a low level of strength in FP1 to a high level of strength in FP2 is $\pi_{1,3}^{1,2}$, and so on.

The transition probability matrix is then computed for all pairs of consecutive FPs. Note that, since H2020 is the last FP included in the EUPRO database, it is not possible to partition it according to the methodology introduced before. Thus, the last transition probability matrix presented in the following Section is the one between FP6 and FP7.

5. Results

The programming budget established by the European Commission to fund the EU FPs has increased over the years. Alongside the provided funds, there has been an evolution from a European policy perspective that has brought an increasing interconnectedness of European regions. As a consequence, European 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.

Figure 1 and Figure 2 highlight this aspect by representing the trend of mean degree and mean strength, respectively, over the different FPs.

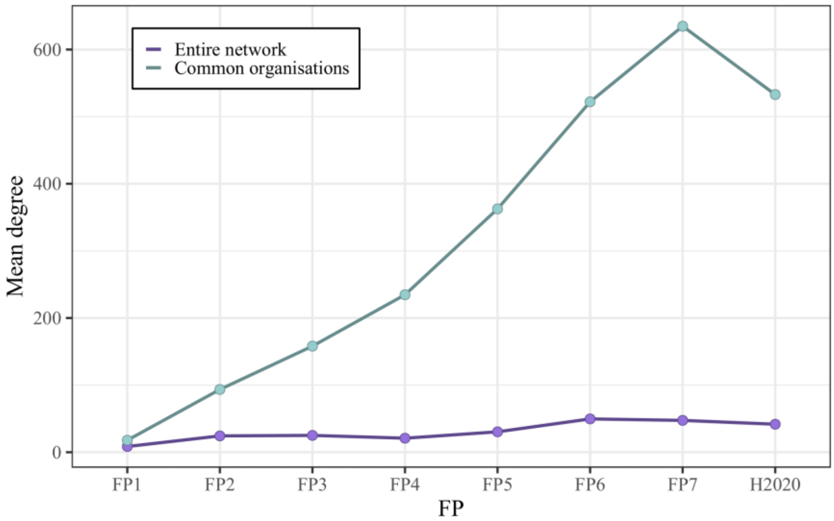

Figure 1: Trend of mean degree: comparison between the entire network and the selected organizations

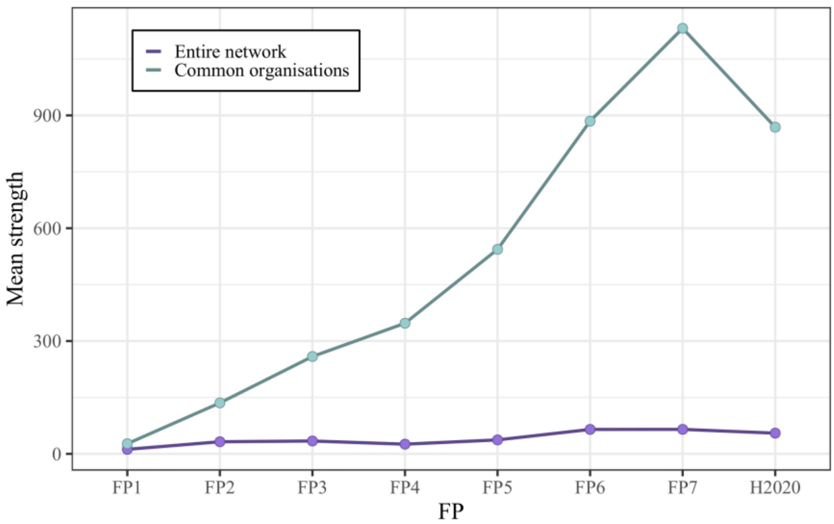

Figure 2: Trend of mean strength: comparison between the entire network and the selected organizations

While participations in H2020 are currently updated, thus affecting results related to the eighth FP, the trend of both metrics slightly increases from FP1 to FP6 within the entire network (except for a decrease in correspondence with FP4), before becoming constant from FP6 to FP7. What is interesting to notice however, is that the trend of both metrics has been dramatically increasing over the FPs for the organizations taking part in all the eight FPs. This result suggests that new incumbents in European research projects tend to collaborate with organizations with previous experience, augmenting the gap between old and new participants over consecutive FPs. This phenomenon is well known in network theory under the name of "preferential attachment" (Newman, 2001), and it has been studied by several authors in the field of scientific collaboration networks (e.g., Wagner and Leydesdorff, 2005; Tomassini and Luthi, 2007; Ferligoj et al., 2015; Zhang et al., 2018). The remarked increase of both metrics also reveals the existence of a certain path dependency, in the sense that participating in previous FPs provides a competitive advantage due to growing experience, competencies, and popularity, as well as other network benefits. Therefore, we expect that the level of centrality of organizations will generally increase from a FP to the subsequent one, i.e., the transitions from low to medium, low to high, and medium to high will be more likely to take place than the transitions from high to low, high to medium, and medium to low.

The estimated transition probability matrices are reported as follows.

$$\pi^{1,2} = \begin{pmatrix} 0.36 & 0.14 & 0.5 \\ 0.15 & 0.09 & 0.76 \\ 0.04 & 0.02 & 0.94 \end{pmatrix};\ \pi^{2,3} = \begin{pmatrix} 0.96 & 0.00 & 0.04 \\ 0.84 & 0.01 & 0.15 \\ 0.24 & 0.01 & 0.75 \end{pmatrix};\ \pi^{3,4} = \begin{pmatrix} 0.48 & 0.01 & 0.51 \\ 0.00 & 0.00 & 1.00 \\ 0.03 & 0.00 & 0.97 \end{pmatrix}$$

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

As expected, the level of strength of the organizations has a higher probability of increasing rather than decreasing in almost all transition matrices. Moreover, it is rather hard for a participant with a high level of strength to move to a less central position in the following FP.

Indeed, the probability of remaining a core institution with a high level of strength is always above 0.71, and in four out of six transition matrices, the value exceeds 0.90. On the other hand, a participant with a low level of strength is more likely to increase its centrality over consecutive FPs, except for the transition from FP2 to FP3 and from FP5 to FP6. While the other four transition matrices have a similar structure, $\pi^{2,3}$ and $\pi^{5,6}$ are quite peculiar. In these two cases, organizations are not particularly boosted by previous participations. Diving deep into these dynamics, we may find a connection with particularly relevant 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, introduced for the first time in 2000, but become fully effective since FP6, marked a substantial shift in European research activities and funding schemes. The promotion of the ERA was indeed aimed to create a common market for science, knowledge, and research at the European level (European Union, 2017). Thus, both events introduced a kind of breakthrough in the functioning of European research programmes and in the way in which funds were assigned, increasing competitivity and openness and partially reducing competitive advantages acquired by previous participants.

6. Discussion and conclusion

In this paper, we propose an innovative approach combining elements from SNA and statistics to map the dynamics of collaborations in projects funded by the EU FPs. Specifically, we take into account all organizations participating in all the first eight FPs, from FP1 to H2020.

We estimate the transition probability matrices (i.e., those matrices whose elements correspond to the probability to move from state i to state j) over consecutive FPs. The final transition probabilities are expressed through a 3×3 matrix, where both rows and columns represent low, medium, and high levels of centralities for two consecutive FPs, respectively.

Levels of centrality are determined based on the computation of the values of strength of the selected sample of participant organizations. We consider strength as the most appropriate centrality measure to proxy core positions in a collaborative network for two main reasons: first, its interpretation is intuitive, deriving straightforwardly from the degree centrality; second, it is more comprehensive than degree centrality, as considering not only the number of partners of an institution among all the projects it takes part in, but weighting more multiple partnerships between the same organizations. Indeed, it is reasonable that the higher the number of collaborations between two entities, the stronger their relationships and their influence within the network. Distinct values of strength are then partitioned for each FP according to an innovative approach, which groups values based on how they evolve in the following FPs. In this way, partitions are determined endogenously rather than exogenously, i.e., based on fixed or quantile thresholds. This procedure represents a contribution itself, being innovative for this field of research, and allowing to obtain unbiased and more precise results in terms of transition probabilities.

Our results support the effectiveness and efficiency of the proposed methodology. Indeed, we find a strong connection between peculiar values obtained in correspondence with two specific transition matrices, i.e., from FP2 to FP3 and from FP5 to FP6, and external phenomena that affected the legal basis and the funding schemes of the FPs during those years. In more detail, we find that almost all transition matrices reveal a path dependency in assuming core positions within the collaborative research networks related to the FPs, and that participants are more likely to maintain or to increase their level of strength over consecutive FPs than shifting to a lower level. However, in the case of the transition matrices from FP2 to FP3 and from FP5 to FP6, the results highlight a higher probability for the organizations to decrease their centrality during the transition from a FP to the subsequent one. We consider this finding a reasonable result since the Treaty of Maastricht first, and the promotion of the ERA then, have been two crucial events determining the openness and the "democratization" of European research funds. Finally, during the course of all FPs, a clear mechanism of "preferential attachment" emerges. These findings provide relevant indications to participants and policy-makers. Our results indeed, clearly show the importance of successfully taking part in European projects to increase the likelihood of obtaining competitive advantages to get funded in future financing programmes. At the policy level instead, some actions are needed to avoid “predatory” behaviors in specific calls for projects, that risk creating a sort of exclusive access to European funds.

It is important to specify that the estimated transition probabilities depend on the dimension of the three partitions. In particular, our approach, which endogenously partitions the distinct values of strength based on how they evolve in a similar way, tends to include a small number of organizations in the medium class compared to low and high categories. This is the reason why values in correspondence of medium levels of centralities are more extreme and less significant. This aspect can be seen as a limitation of our analysis.

In the future, we want to investigate the transition probabilities moving more than one-step forward in terms of FPs, focusing on those organizations that kept a high level of centrality over all the FPs. Furthermore, we aim to replicate the analysis introducing some variations and constraints to the model. We are also interested in applying the methodology to the dynamic study of other collaborative research networks to see if our approach confirms the ability to detect the impact of external events. This aspect has indeed relevant policy implications providing indications to European institutions about the functioning of research programmes as well as motivation to participants to increase their centrality in order to gain competitive advantages.

References

Amoroso S, Coad A, 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.

Arranz J, de Arroyabe JCF (2012), Can innovation network projects result in efficient performance?, Technological Forecasting & Social Change 79:485-497.

Balland PA, Boschma R, Ravet J (2019), Network dynamics in collaborative research in the EU, 2013-2017, European Planning Studies 27(9):1811-1837.

Borgatti SP, Halgin DS (2011), Analyzing affiliation networks, In: Carrington P, Scott J (Eds.), The Sage handbook of social network analysis, SAGE publications.

Börner K, Sanyal S, Vespignani A (2007), Network science, Annu Rev Inf Sci Technol 41(1):537–607.

Breschi S, Cusmano L (2004), Unveiling the texture of a European Research Area: emergence of oligarchic networks under EU Framework Programmes, International Journal of Technology Management 27(8):747-772.

Cavallaro M, Lepori B (2021), Institutional barriers to participation in EU framework programs: contrasting the Swiss and UK cases, Scientometrics 126(2):1311–1328.

Cerqueti R, Falbo P, Pelizzari C (2017), Relevant states and memory in Markov chain bootstrapping and simulation, European Journal of Operational Research 256:166-177.

Cinelli M, Ferraro G, 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.

Commission of the European Communities (2002), The European Research Area. Providing new momentum. 565 final, Brussels, 16 October.

Commission of the European Communities (2007), Green paper "The European Research Area: New perspective". 161 final, Brussels, 4 April.

Crespo J, Suire R, Vicente J (2016), Network structural properties for cluster long-run dynamics: Evidence from collaborative R&D networks in the European mobile phone industry, Industrial and Corporate Change 25(2):261–282.

Enger SG, Castellacci F (2016), Who gets Horizon 2020 research grants? Propensity to apply and probability to succeed in a two- step analysis, Scientometrics 109:1611-1638.

Enger SG (2018), Closed clubs: Network centrality and participation in Horizon 2020, Science and Public Policy 45(6):884-896.

European Union (2017), EU framework programmes for research and innovation: Evolution and key data from FP1 to Horizon 2020 in view of FP9, September 2017.

Ferligoj A, Kronegger L, Mali F, Snijders TA, Doreian P (2015), Scientific collaboration dynamics in a national scientific system, Scientometrics 104(3):985-1012.

Heller-Schuh B, Barber M, Bilalli Shkodra X, Scherngell T, Zahradnik G (2020), Documentation of crisis datasets: Eupro.

Heringa PW, Hessels LK, 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.

Hoekman J, Scherngell T, Frenken K, Tijssen R (2013), Acquisition of European research funds and its effect on international scientific collaboration, Journal of Economic Geography 13:23-52.

Hu W, Gong Z, U LH, Guo J (2015), Identifying influential user communities on the social network, Enterprise Information Systems 9(7):709–724.

Lepori B, Veglio V, Heller-Schuh B, Scherngell T, Barber M (2015), Participations to European Framework Programs of higher education institutions and their association with organizational characteristics, Scientometrics 105:2149-2178.

Mokhtarzadeh NG, Mahdiraji HA, Jafarpanah I, Jafari-Sadeghi V, Cardinali S (2020), Investigating the impact of networking capability on firm innovation performance: using the resource-action-performance framework, Journal of Intellectual Capital 21(6):1009-1034.

Newman MEJ (2001), Clustering and preferential attachment in growing networks, Physical review E 64(2): 025102.

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

Roediger-Schluga T, Barber MJ (2008), R&D collaboration networks in the European Framework Programmes: data processing, network construction and selected results, Internation Journal of Foresight and Innovation Policy 4(3-4):321-347.

Scherngell T, Barber MJ (2011), Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU Framework Programme, The Annals of Regional Science 46:247-266.

Scherngell T, 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.

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

Tomassini M, Luthi L (2007), Empirical analysis of the evolution of a scientific collaboration network, Physica A: Statistical Mechanics and its Applications 385(2):750-764.

Uhlbach WH, Balland PA, Scherngell T (2017), R&D Policy and Technological Trajectories of Regions: Evidence from the EU Framework Programmes, SSRN Electronic Journal pp 1–21.

Wagner CS, Leydesdorff L (2005), Network structure, self-organization, and the growth of international collaboration in science, Research policy 34(10):1608-1618.

Wanzenböck I, Neuländtner M, Scherngell T (2020), Impacts of EU funded R&D networks on the generation of key enabling technologies: Empirical evidence from a regional perspective, Papers in Regional Science 99(1):3–24.

Zhang C, Bu Y, Ding Y, Xu J (2018), Understanding scientific collaboration: Homophily, transitivity, and preferential attachment, Journal of the Association for Information Science and Technology 69(1): 72-86.