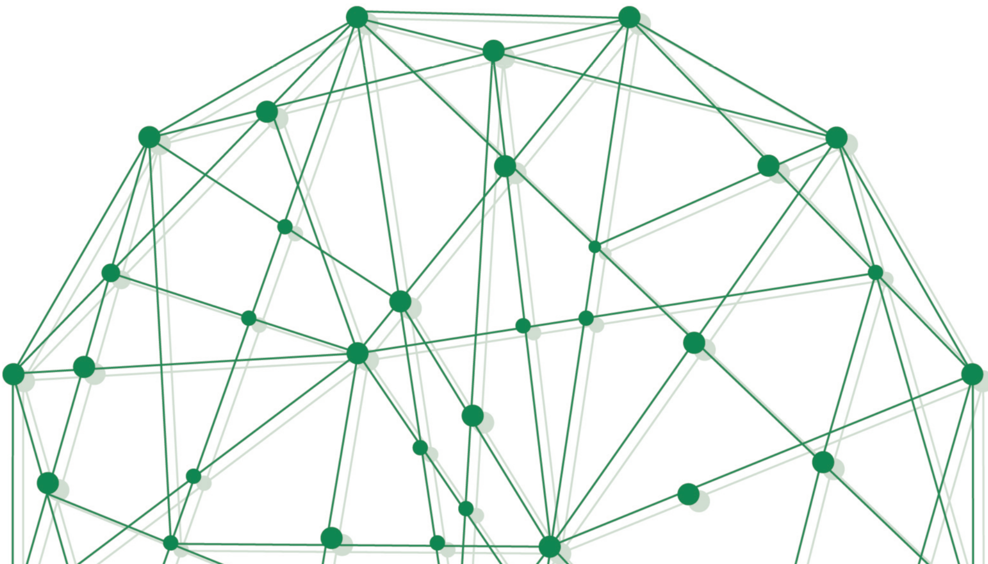
Edited by Giuseppe Giordano, Marialuisa Restaino, Andrea Salvini

METHODS AND APPLICATIONS IN SOCIAL NETWORKS ANALYSIS

Evidence from Collaborative, Governance, Historical and Mobility Networks



COMPUTATIONAL SOCIAL SCIENCE

FrancoAngeli
OPEN ACCESS

COMPUTATIONAL SOCIAL SCIENCE

La collana accoglie contributi di carattere interdisciplinare relativi al dibattito sul campo derivate dall'applicazione di metodi innovativi di ricerca e pratiche di uso dei Big Data, con un'attenzione particolare alle tematiche epistemologiche, metodologiche e politiche di gestione dei contenuti digitali.

Secondo la letteratura internazionale è possibile definire la scienza sociale computazionale come una disciplina che sfrutta la capacità di vasti set di Big Data per analizzare le interazioni umane al fine di definire prospettive qualitativamente nuove sul comportamento collettivo, in un approccio interdisciplinare che comprende sociologia, statistica, informatica, psicologia, diritto, matematica e fisica teorica.

La ricerca sociale computazionale, basandosi sull'analisi delle tracce digitali delle attività online, l'analisi dei network sociali, le fonti aperte digitali, la simulazione sociale attraverso modelli computazionali, rappresenta uno strumento proficuo per l'analisi del mutamento sociale. In tale direzione essa ha già prodotto, negli ultimi dieci anni, moltissimi contributi che confermano la rivoluzione metodologica in atto.

All'interno di questa cornice e in considerazione della crescente consapevolezza della comunità scientifica internazionale di quanto la ricerca sociale debba passare necessariamente per un utilizzo attivo delle tecnologie dell'informazione, la collana ha quindi come obiettivo principale la costituzione di uno spazio di discussione epistemologica, ontologica e metodologica interdisciplinare nel quale poter raccogliere, valutare e catalogare i contributi specifici dell'analisi computazionale.

I volumi pubblicati, in lingua italiana o inglese, sono sottoposti alla valutazione anonima di almeno due referees esperti nei settori scientifico-disciplinari della matematica, della sociologia, della statistica, della fisica teorica, del diritto, dell'informatica e della psicologia.

Edited by Giuseppe Giordano,
Marialuisa Restaino, Andrea Salvini

METHODS AND APPLICATIONS IN SOCIAL NETWORKS ANALYSIS

Evidence from Collaborative, Governance, Historical and Mobility Networks

COMPUTATIONAL SOCIAL SCIENCE

FrancoAngeli
OPEN ACCESS

This volume is published with the contribution of the Department of Political Sciences (University of Pisa) and the Department of Political and Social Studies (University of Salerno).

This volume is published in open access format, i.e. the file of the entire work can be freely downloaded from the FrancoAngeli Open Access platform (http://bit.ly/francoangeli-oa).
On the FrancoAngeli Open Access platform, it is possible to publish articles and monographs, according to ethical and quality standards while ensuring open access to the content itself. It guarantees the preservation in the major international OA archives and repositories. Through the integration with its entire catalog of publications and series, FrancoAngeli also maximizes visibility, user accessibility and impact for the author.

Read more:
http://www.francoangeli.it/come_pubblicare/pubblicare_19.asp

Index

Preface

by Giuseppe Giordano, Andrea Salvini, Marialuisa Restaino

ARS Conferences and the development of Social Network Analysis

This book contains some of the contributions presented at the ARS'19 Conference held in Vietri sul Mare (Salerno, Italy) in October, 29-31 2019; it was the seventh of a biennial meetings series started in 2007 with the aim to promote relevant results and the most recent methodological developments in Social Network Analysis.

ARS stands for "Analisi delle reti sociali" in the Italian language: it is, therefore, a Conference that since its first edition has brought together scholars from various disciplines within the social sciences, with the aim of sharing their research and contributing to the dissemination and advancement of Social Network Analysis (SNA) on a theoretical, methodological and applicative level.

The success of the ARS Conferences has grown along with the affirmation of the network perspective in every dimension of scientific knowledge; when the first ARS Conference was held in 2007, SNA had not yet taken on the characteristics of a widespread and shared paradigm, especially in the social sciences – although it had already been consolidated internationally for at least three decades. The fundamental criticisms of this paradigm at that time were still affected by the deep division that had been created in previous decades around the theoretical dichotomies of reductionism/determinism and structure/action, and by the contrast between quantitative and qualitative methods. In this framework, SNA was considered by many as a partial and determinist perspective because it was exclusively concerned with the structural dimension of phenomena, and spoiled by an obsessive use of quantitative technicalities. This criticism could not have been more incorrect.

Over the last twenty years, the SNA perspective has proven the ability to develop a significant breadth of theoretical and methodological elaboration, through the contribution of a growing number of scholars and the multiplication of empirical applications in all fields of social sciences and not only social sciences. One of the disciplinary areas in which this development has occurred, among others, is certainly that of computational social science, by virtue of the growing investment in the field of online social networks and the role of information technologies in the production of scientific knowledge. The fact that this volume is published in a series dedicated to computational social sciences, demonstrates the successful convergence between the two disciplines.

The growing awareness of the complex nature of social phenomena enforced the usefulness of the network perspective as a wealth of theoretical and methodological tools capable of appropriately penetrating within the dimensions of that complexity. It is no coincidence that the 2019 edition of ARS has been mainly addressed to the study of *Multilayer, Multilevel and Multimode Networks*. Obviously, the methodological and applicative topics were very large and varied. It should also be highlighted that the ARS conferences have taken on an international vocation since the first editions, as demonstrated by the presence of scholars from all over the world and of different disciplinary origins.

Finally, though a conference event is a place where specialized scholars meet together to give evidence of their recent researches, the ARSs conferences have always paid great attention to younger scholars and to the actual impact that academic research can furnish to the society, in a broader sense.

These aspects confirm ARS as one of the most significant appointments in the panorama of scientific events dedicated to SNA – also thanks to the effort of the group of organizers to promote the dissemination of the most significant content emerging from the ARS Conferences through specific international publication. Specifically, retracing the past editions of the Conference, we recall the related publications:

- ARS’17 – May 15-17, 2017, “Challenges in Social network research”, Special issues in the *Italian Journal of Applied Statistics*, 30, 2018;
- ARS’15 – April 29-30, 2015, “Large Networks and Big Data: New Methodological Challenges”, Special issue in *Social Network Analysis and Mining*, 6, 2016;
- ARS’13 – June 20-22, 2013 “Networks in Space and Time: models, data collection and applications”, Special issue in *Network Science*, 3, 2015;
- ARS’11 – June 23-25, 2011 “Collaboration Networks and Knowledge Diffusion: Theory, data and methods”, Special issue in *Industry and Innovation*, 20, 2013;

– ARS’09 – July 13-14, 2009 “Social Network Analysis: Models and Methods for Relational Data”, Special Issues Advances in *Data Analysis and Classification*, 5, 2011 – *Quality & Quantity*, 45, 2011;
– ARS’07 – November 30, 2007 “Per conoscere uno strumento – Uno strumento per conoscere”, Special issue AA.VV., *Analisi delle Reti Sociali: per conoscere uno strumento, uno strumento per conoscere*, vol. 17, Rubbettino, Soveria Mannelli (in Italian).

In line with this tradition, the Scientific Committee of ARS’19 decided to encourage some scholars to present their work in an extended version addressed to a wider area of possible readers.

This book

The book hosts eleven contributions that within a sound theoretical ground, present different aspects of the applicative fields where the Social Network Analysis can contribute to explore, interpret and predict interaction and social behavior between actors. The formalism and the technical terminology were maintained, when necessary, only for sake of scientific accuracy while the reference to real case studies may help the reader to discover the potentiality and the interpretative power offered by the Social Network Analysis.

The text is mainly articulated into a first part that is more strictly methodological, and a second part that contains essays of a more substantive nature. However, as will be clear during the reading, both dimensions support each other, given the close connection that exists, within the SNA, between the substantive working hypotheses and the choice of consistent analytical techniques.

The chapter edited by Bothorel *et al.* is dedicated to community detection methods in complex networks. The paper proposes an original methodology for the selection of the algorithm, based on hub dominance and transitivity, according to which it is possible to identify communities with appropriate characteristics with respect to the observed network structure. The application of this methodology to a case study (the crowdfunding platform “Ulule”), shows the effectiveness of the proposal and confirms how the concept of homophilia plays an essential role in the structuring of online groups.

The next chapter, by Sepulvado *et al.*, shifts the focus to the temporal dimension in which network ties are formed, evolve and persist, introducing the notions of dyadic similarity trajectory and dyadic similarity trajectory cluster. The two notions allow to predict the temporal dynamics of a network through the analysis of the way in which the patterns of similarity between dyads change over time, and eventually transform or decay. Further developments of this approach can involve the use of other endogenous measures to the network, in addition to that of homophilia, such as that of centrality, in order to obtain more refined studies of the link between the trajectories of the cluster and actors' characteristics in a temporal perspective.

The chapter edited by Piazza and Vasudevan, is part of a consolidated tradition of studies, and shows the usefulness of the use of the network autocorrelation model in assessing the mechanisms of social influence based on the presence of interdependent individuals embedded in the networks. The model is applied within a survey conducted in a university context, aimed at verifying whether the relationships between students positively influence their academic performance. The essay focuses on the construction of relational variables and those used to verify the working hypothesis, keeping under control possible effects on academic performance that may arise from other factors. Among the outcomes of the research it is worth highlighting how the processes of interaction between students in university contexts clearly influence their outcomes in terms of outcome and academic satisfaction.

Riccardo and Psaroudakis, in the chapter that closes the first part of the volume, apply the techniques of the multilevel exponential random graph model (MERGM) to study the structural configuration of cross-level interactions within two Third Sector inter-organizational networks operating in Southern Italy. The multilevel models in the context of the proposed research, have allowed to test hypotheses on how and how much the presence of links to the lower-lever of social actors (such as organizations) depends on the links between actors at higher levels (such as those involving organizations and their activities). The results of the elaborations have led the authors to affirm that the system of interdependencies that is created through participation in social interventions at the community level, strengthens collaborative networks making the outcomes more effective in terms of welfare for the territorial communities served.

Part II of Volume opens with an essay by Marotta et al., that is in line of continuity with the previous one, since it has as its object the study of the network structure constituted by collaborative interactions between organizations that carry out social policy interventions on some small municipalities in Southern Italy. The descriptive analysis of the network parameters is aimed at offering useful indications to make the network governance more effective and efficient and to identify possible directions to improve qualitatively and quantitatively the relationships of collaboration between collective actors within the welfare community. Of particular interest is the evidence that in small municipalities the mutual knowledge between social actors (individual and collective) is still an important condition for the development of meaningful collaborative relationships.

The chapter edited by Trebitz and Shrestha, focus on drivers of the core-periphery structures observed in 18 water governance networks related to fisheries and/or water quality in five large reservoir basins of the greater Columbia River Basin. Adopting the framework of resource dependency theory and a mixed methodological approach, the authors argue that the governance under study is characterized by a core-periphery network structure, in which the centrality of the actors is driven by the roles they act in the network defined not only by specific institutional characteristics, but also by the resources they possess and the informal behaviors that are acted in the network.

The seventh chapter, edited by Shrestha, uses the techniques of comparison group evaluation and bipartite exponential random graph models in the case of network building of the rural communities with organizations in Nepal after the end of the project assistance. The results of the analysis work show that there is no difference between the communities with and without the assistance, which raises questions about the extent or the duration of the project support needed to make a difference. This essay demonstrates the importance of the application of social network analysis in the evaluation of intervention projects – an aspect scarcely present in the literature – and how the analysis of the communities' networks with organizations could serve as a useful diagnostic tool for community development professionals engaged in improving community capacity and network building.

The chapter of Russo et al, projects us into the contemporary atmospheres of the spread of the Covid-19 contagion and proposes a contribution that analyses the dissemination of information in the Twittersphere during the COVID-19 health emergency in Italy, with a specific focus on the role of social actors in the social media of Twitter and how they are connected to each other. The exploratory analysis aims to provide a first photograph of the characterization of the flow of information related to the main production and dissemination agents of content on Twitter. Applying techniques of semantic network analyses, the authors show how, in Covid's time, the stage of online communication has been occupied mainly by Bloggers, Media, Journalists, and satirical commentators, who show to perform mainly political polarization function.

The ninth chapter, by Lorandini and Odella, introduces us to the fascinating dimension of social network approaches to historical data; the study presented aims at investigating the structure of lending relations in pre-modern economy, in particular focusing on the network established by a merchant family along approx. forty years (1747-1786) and scrutinizes hypothesis concerning the social and economic mechanisms of early credit markets. The circumstance for which the hypothesis relating to the structural configuration of lending relationships has been partially fulfilled, shows how the network approach applied to financial relationships – in a historical perspective – must be accompanied by careful consideration of the socio-economic contexts in which differentiated approaches to credit have occurred.

The chapter edited by Del Forno and Di Gregorio, proposes a study on ego-networks of innovative startups "with social vocation" in Piedmont and Campania, with the aim of recognizing the signs of possible integration between the sphere of public authorities, the business community, and the organized civil society. Adopting a mixed-methods approach the authors were able to compare the characteristics of ego-networks of innovative enterprises operating in the two Italian regions, confirming the diversity of the welfare structure at local level.

The essay that concludes the book, by Leoni and De Angelis, proposes a study in which the network perspective is applied to the dynamics of mobility of Erasmus students, both incoming and outgoing from Italian universities, in the academic year 2013/2014.

The results of the analysis that describes the flows of students show various situations of heterogeneity between groups of states in Europe, which are often expressed in evident asymmetries both in the dimension of the balance between incoming and outgoing students, and in the disciplines of study – which see the preeminence of human, social and engineering sciences, to the detriment of natural sciences and ICT.

The contributions presented here are only a small cross-section of the empirical and theoretical wealth that is provided by the Social Network Analysis, and the extraordinary variety of areas of exploration that can benefit from its analytical tools. Sometimes, the large number of research experiences do not find an adequate space to build meaningful syntheses of the progress achieved. Therefore, there is an increasing need for events and tools that allow to reduce the fragmentation of these experiences, and to foster a broad interdisciplinary reflection on the development and possibilities of the SNA perspective.

There is also the need to guarantee the younger generations of scholars a wider and more reliable access to training and sharing opportunities with respect to the development of methods and techniques of social network analysis. We are convinced that volumes such as the present one and events such as those guaranteed by the ARS Conferences can contribute to the pursuit of these objectives.

Acknowledgements

The Volume Editors wish to thank the Directors of the FrancoAngeli Series Computational Social Science, Mara Maretti (University of Chieti-Pescara, Italy) and Lara Fontanella (University of Chieti-Pescara, Italy), and all the anonymous reviewers who contribute to improve the whole publication, without their contributions this volume would not have been realized.

We finally thank Maria Prosperina Vitale for her support and advices.

This book was financially supported by the Department of Political and Social Studies, University of Salerno and the Department of Political Science, University of Pisa.

Part I
Research Methods in SNA

1. How to choose community detection methods in complex networks to study cooperation and successful organizations. The case study of Ulule crowdfunding platform

by Cécile Bothorel[*], Laurent Brisson[**], Inna Lyubareva[***]

1. Introduction

In recent years socio-economic research on online groups and communities often proposes to extend the traditional approach and to encompass social network analysis modelling relationships by edges in the graph. The joint analysis of the two types of data-socio-economic and network structure – makes possible to provide important insights on the group functioning and to reveal properties of a social network that are not immediately obvious, e.g. the existence of sub-networks or communities operating within the global network.

For instance, in the context of crowdfunding platforms, many studies focus on the directly observable interactions among the participants of individual projects and show their role for the success of fundraising campaigns (Kuppuswamy & Bayus, 2018; Agrawal, Catalini & Goldfarb, 2010; Zheng, Wu & Xu, 2014). However, few empirical studies focus on relational structures which are *not explicitly* stated in the context of crowdfunding platform (Inbar & Barzilay, 2014) and question whether relational circles can go beyond individual projects, by broadening the initial social capital of projects' leaders, and form a platform-level cross-project social network. Does the participation of the members of this cross-project network guarantee higher success rates of crowdfunding campaigns? These questions have strong man-

[*] IMT Atlantique, LUSSI Department, Lab-STICC CNRS UMR 6285, cecile.bothorel@imt-atlantique.fr.

[**] IMT Atlantique, LUSSI Department, Lab-STICC CNRS UMR 6285, laurent.brisson@imt-atlantique.fr.

[***] IMT Atlantique, LUSSI Department, LEGO EA2652, inna.lyubareva@imt-atlantique.fr.

agerial and economic implications for the platforms: *Could this networking lead to the formation of a core group with an intense participation? Should the platforms further develop the mechanisms for cross-project networking? How should project leaders surround themselves to give themselves every chance of success?* The study of these questions necessitates an appropriate methodology based on the detection of *implicit* communities. Therefore, the question arises of how to choose an appropriate community detection algorithm, relevant to the particular context.

As a matter of fact, many community detection methods exist. They have different ways to divide a network into multiple subsets of densely connected nodes, and hence result in different community structures. They may rely on different notions of community, and even with the same notion (e.g. Newman defines a community as a "group of vertices with a higher-than-average density of edges connecting them" (Newman, 2006), algorithms may optimize different objective functions or use different heuristic to get efficient implementations. The authors of the methods compare their efficiency in terms of computational properties, such as complexity, as well as in terms of validation metrics: modularity, rand index, conductance, etc. (Fortunato & Hric, 2016), describing resulting partitions quality. However, as demonstrated in some comparative surveys (Dao, Bothorel & Lenca, 2020; Fortunato & Hric, 2016; Jebabli *et al.*, 2018), choosing a community detection method is still a problematic task, especially if we are not aware of the underlying mechanisms.

Some recent papers have attempted to provide guidance on algorithm selection using criteria such as the mixing parameter of a network, computation time, or overlap with a simulated community structure (Dao, Bothorel & Lenca, 2020; Zheng, Wu & Xu, 2014). Others provide more intuitive criteria such as the size of the communities (Dao, Bothorel & Lenca, 2019) or descriptive network-oriented metrics on structural patterns within communities (Dao, Bothorel & Lenca, 2017, 2018). N. Smith *et al.* argue that such criteria should not be applied alone, but in conjunction with business-oriented objectives: the "best" method depends on the context, on the research question, i.e. on how the communities will be used (Smith *et al.*, 2020).

The current paper addresses this problem. It proposes an original methodology to guide practitioners in their choice of methods in connection to a specific research question. We first pre-select some candidate methods, and qualify their results, i.e. the resulting partitions, through qualitative characteristics. Then, final users may compare the few pre-selected partitions with comprehensive measures and finally select the most relevant one regarding the research question.

The case of the Ulule crowdfunding platform is used as example. In this case study, we suggest that some *non-directly observable* communities may exist within the cross-project social network. However, as in many exploratory studies, we have no preliminary information about their number or structures. Our goal is to discover these communities, to describe their internal organizations and to explore whether there is a relationship between particular communities' organization and the success of fundraising campaigns. This last element is the central research question of the case study.

This paper is organized as follows: In Section 2, we introduce the Ulule platform and its cross-project social network. In Section 3, we present a methodology which permit us to select, among 11 methods, three most relevant and convergent algorithms to discover the communities. In Section 4, by introducing additional socio-economic attributes of communities, we choose the most suitable method and discuss the relative success of different communities for fundraising campaigns. Last Section concludes.

2. Ulule crowdfunding platform

2.1. Ulule platform and its network

Crowdfunding represents a model of participatory financing, which has been used by an increasing number of companies, associations and individuals since the early 2010s. Its principle consists for a project "pitched" by its creator, to collect money from a large public. Thanks to numerous interaction mechanisms between participants (comments, news, promotion systems, etc.), these platforms play the role of facilitators of social capital.

To better understand the networking role of crowdfunding platforms, we use empirical data for the period 2010-2016 from the Ulule platform – one of the main crowdfunding platforms in France and Europe, which pays particular attention to the strengthening of its proper platform's community[1]. After data cleaning, we keep 19,544 projects, of which 11,900 were successfully funded, and 7,644 failed. They globally attracted 876,758 contributors, who contributed a total of 47.75 million euros.

[1] https://www.ulule.com/about/ulule/.

2.2. Graph of co-contributions

While 99.7% of Ulule contributors are one-time funders, 0.3% of them are not only very active by contributing to more than three different projects, but they regularly "meet", in each fundraising campaign, the same participants also contributing to these projects. We suggest that these active contributors, involved in a least 3 projects with at least 1 other contributor, are candidates to form the cross-project platform's network. On average, they contribute to 14 different Ulule projects, for an average total amount of over 600 euros per contributor. With these active Ululers, we define a non-oriented graph of co-contributions, in which each edge (u,v) means that the users u and v have contributed to 3 or more projects together. There are 469 connected components. All of them contain less than 10 nodes, except the largest one. With its 2,081 nodes and 4,749 edges, this giant component proves that the social network, transverse to individual projects, really exists at the Ulule platform.

The members of the cross-project network share projects with 4.56 other members (average degree). We find a power law distribution of degrees (the estimated exponential coefficient is 2.27) which is a common property in online social networks. 25% of the members have a degree greater than 4, the maximum degree being 199. There are therefore contributors who co-finance projects with really important number of different contributors: 24 of them have more than 50 neighbors in the graph.

We notice that the distribution of the Ulule's thematic categories according to whether or not projects are financed by the members of the graph reveals interesting information. The improvement in the success rate of campaigns of projects belonging to the graph is observed for all categories (28.6% overall improvement), especially for Games, Comics, Technology and Publishing. Indeed, these domains have a strong social component in the production and consumption of goods. These findings confirm our interest in studying communities of the active Ululers.

We find out that members of the obtained social graph vary a lot in terms of their social activity (node's degree, clustering coefficient and centralities) and contribution behavior (number of projects funded, average amount of contributions, specialization rate which quantify the variety of thematic categories addressed (projects' category) and their similarities with the neighboring nodes).

Multiple correspondence analysis on these attributes, followed by a hierarchical bottom-up classification, leads to 5 clusters of Ulule's active contributors. A detailed version is available (Lyubareva *et al.*, 2020), but for space reasons, we sum up the 5 profiles as follows:

- the Sponsors (18) are the Ululers with a high degree and betweenness centrality. They may be considered as facilitors. Involved in 140 projects (in average), they are very active on various topics (not specialized);
- the Followers (653) are the users who arrive late, during the second half of the campaigns. They prefer very big projects (which average objective is more than 17k euros). Exhibiting a high closeness centrality, they frequently co-contribute with other contributors, in particular the Sponsors;
- the Precursors (538) are characterized by an early arrival in projects and especially before all their neighbors;
- the Collaborative Specialists (368), very highly thematically specialized, have also a very high clustering coefficient indicating a strong cohesion of links between neighbors. This may show a strong solidarity between some Ululers whose financing decisions are often collective, even if their fields of specialization are not necessarily identical. They don't contribute to a very large number of projects and are not attracted by the size of the projects;
- the Specialists (504) are not high contributors as well in terms of number of projects, nor in volume of contributions. They are not particularly highly connected as well, but they are however passionate about very specific themes, the same themes than their neighbors, demonstrating therefore social homophily.

These findings will be mobilized in the Section 4 to refine the final choice of the community detection method in connection with the case study research question.

3. A methodology to choose community detection methods

There are many approaches to perform community detection based on different paradigms, including cut, internal density clustering, stochastic equivalence, flow models, etc. (Fortunato and Hric, 2016). The purpose is not to provide an exhaustive overview here. We refer the reader to surveys like (Fortunato & Hric, 2016; Schaub *et al.*, 2017) to get details about these different approaches. In this work we focus on well-known methods, apply them to our case- specific graph and show how different the partitions produced can be, making the choice of a method non obvious.

We carefully kept a large variety of approaches as summarized in Table 1. While Edge betweenness is based on edge centrality detection in order to split networks into several communities, Louvain and Fast greedy optimize modularity by iteratively folding nodes into meta-nodes. Spectral method is also based on modularity, but identifies the community structure by finding leading eigenvectors corresponding to largest eigenvalues of a modularity matrix. Some approaches are based on a dynamic distance: for Walktrap, if two nodes are in the same community, the probability that a random walker will move from one to another in only a few movements is very high (notion of trap) and consequently the distance is low. Conclude combines a similar random walk-based distance to agglomerate nodes, and the local optimization of modularity inspired by the Louvain method selects iteratively the best agglomeration. Infomap relies on finding a configuration that maximizes the compression of random walks represented by an encoded binary sequence. Inspired by epidemic spreading mechanisms, a more basic and direct distance is used by the Label propagation and its variant SLPA, where a node should belong to the same community as most of its immediate neighbors. In the same vein, concepts have also been borrowed from theoretical physics with the Spin glass model which may be seen as an alternative to modularity maximization; the idea is to consider nodes as spin states and to minimize the energy of the configuration of spins to reach a stable state. Finally we can cite a statistical inference approach, DCSBM which uses stochastic blockmodels to infer the likeliness that a given observed network (and its latent block structure) is generated from a compatible model, and then suggest the most likely set of model parameters. Methods are often compared with classic quantitative measures from information theory domain, such as NMI, the Normalized Mutual Information (Chakraborty *et al.*, 2017), which evaluates their agreement to arrange nodes into similar clusters. Recent studies introduced more intuitive, and very simple quantitative metrics such as the size of the communities (Dao, Bothorel & Lenca, 2019) where methods are similar if they produce similar community size distribution. In (Dao, Bothorel & Lenca, 2017), the authors propose metrics dedicated to graph analysts. They describe the communities with structural measures, in order to quality the communities from a topological point of view. Each community is qualified with well-known organizational patterns, such as star-based structures, cliques (this will be detailed later, in Figure 3).

In this paper, we propose to practitioners a methodology based on both these quantitative and qualitative metrics. We show how to use them in order to compare the methods; through the case study from Ulule, and its specific business question – which communities lead to successful fundraising –, we show how to select the most relevant method:

- the initial step is to run community detection algorithms. We used 11 methods described in Table 1 and 11 partitions (where each node belongs to only one community) were obtained;
- step 1 aims at the choice of a subset of methods. We compare the partitions with validation metrics. Since we have no preliminary knowledge on the

platform's communities, we are looking for the methods making consensus findings in order to build a robust foundation for subsequent steps;
- as proposed in the previous literature, Step 2 brings an additional intuitive criterion, the size of communities, to refine the subset of consensual methods;
- step 3 characterizes the consensual partitions with qualitative measures that are relevant for the current business-oriented problem. As our case study focuses on organizational patterns involving nodes, we will use bivariate maps based on graph structural indicators, such as the hub dominance and the clustering coefficient;
- the last step introduces specific "business" indicators related to the current problem to finally make the choice decision (section 4).

3.1. Initiate the selection with consensual partitions

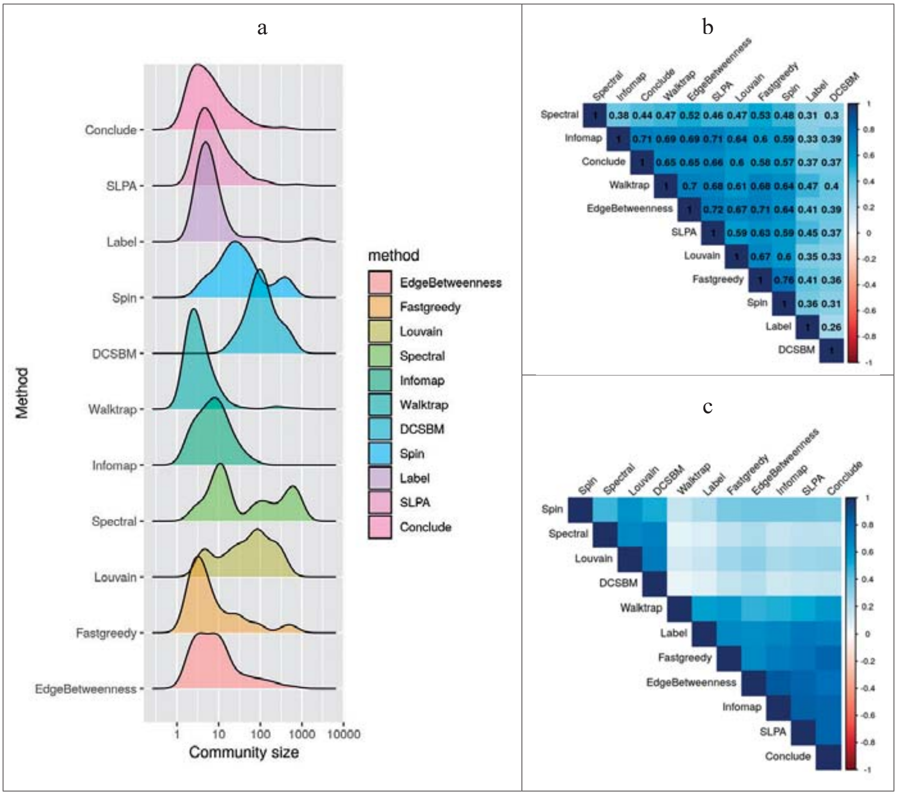
After the initial step, that is run the 11 methods with their default parameters), we obtain 11 partitions. In the first selection step, we compare how nodes are arranged into clusters: we compute the Normalized Mutual Information (NMI) often used in community detection because it allows the comparison of partitions even where nodes are assigned to a different number of clusters. We apply its normalized variant with values in [-1, 1] which is popular in the field of community analysis (Chakraborty *et al.*, 2017). The Figure 1b shows how our 11 partitions are astonishingly similar from a NMI point of view where all scores are positive, ranging from 0.26 to 0.76, with a large majority above 0.5. Eight among them demonstrate scores greater than 0.6 when compared to each other, with a very consensual group consisting of four methods: Edge Betweenness, SLPA, Fast greedy and Walktrap with scores higher than 0.7. There are three slightly different methods: Spectral, Label Propagation and Spin Glass that produce more specific partitions, which are all different from each other. While this difference may be easily explained by the fact that Spectral and Spin Glass implement inherently distinctive mechanisms, the results of Label Propagation, being a variant of SLPA, are quite surprising.

As an intermediate result, on the basis of the NMI scores, Edge Betweenness, SLPA, Fast greedy, and Walktrap converge on their clustering task. To put it differently, knowing a random node's affiliation in Edge Betweenness partition, in our example, gives a high probability to successfully deduce its membership in the 3 other partitions. Louvain is also quite close to this group and as it is frequently used, we don't want to discard it right now.

Tab. 1 – A summary of community detection methods used to study community structure

Method	*Approach*	*Reference*
Louvain	Multilevel modularity	Blondel *et al.* (2008)
Fast greedy	Modularity optimization	Clauset *et al.* (2004)
Spectral	Vector partitioning	Newman & Girvan (2004)
Spin glass	Energy model	Reichardt & Bornholdt (2006)
DCSBM	Stochastic blockmodels	Karrer & Newman (2011)
Walktrap	Dynamic distance	Pons & Latapy (2005)
Conclude	Dynamic distance	Meo *et al.* (2014)
Edge Betweenness	Edge centrality detection	Girvan & Newman (2002)
Infomap	Information compression	Rosvall & Bergstrom (2008)
Label propagation	Topological closeness	Raghavan *et al.* (2007)
SLPA	Topological closeness	Xie & Szymanski (2011)

Fig. 1 – The similarity between community detection methods in terms of (a) Community size distributions, (b) NMI, (c) size fitting quality



These 5 methods reach a quite good consensus. We argue that in exploratory studies, as the current one, where practitioner has no a priori knowledge about communities that she wants to analyze, it is important to identify such robust clustering, demonstrating an agreement between different methods. This step is determinant for the further exploration of the research question.

3.2. The size of communities as qualitative choice

We propose then to complete the NMI analysis by adding information about the size of the discovered communities (Fig. 1a). One can notice two types of partitions. Some communities are large with tens or sometimes thousands of members. Other partitions, on the contrary, exhibit small (2 or 3 members) to medium-sized communities (around 10-20 nodes). These two classes of methods are depicted in Figure 1c, where the similarity score compares the distribution of sizes of communities (score introduced in Dao, Bothorel & Lenca, 2018).

Regarding large communities, Louvain, Spin glass and Spectral produce an interesting variety of large to very large communities, whereas DCSBM only produces huge ones. We observe the same disparity when we focus on partitions with smaller groups. SLPA and Walktrap produce a lot of communities with approximately 10 nodes, without notable variety of sizes. On the contrary, Fast greedy's and Edge Betweenness partitions have a more flat profile in Figure 1a: most of their communities are small to medium, but some of them have also more than 100 nodes.

Therefore the consensual methods – Louvain, Edge Betweenness, SLPA, Fast greedy, and Walktrap – produce different distributions of sizes of communities, we have to choose whether we give priority to small-medium communities or to large ones. This example demonstrates that whereas in some contexts, the size criteria may be enough to make a choice, such information is not always sufficient. Taking into account that our research question focuses on the communities' forms and their efficiency for the fundraising campaigns, we add topological indicators in order to differentiate the partitions from the organizational perspective.

3.3. Structural classification of consensual partitions

In order to characterize organizational patterns within communities, which is our business-oriented objective here in the case study, we propose the use of structural measures applied to communities such as internal link density, average centrality of nodes, average degree. Such indicators are interesting to be combined in bivariate map (Dao, Bothorel & Lenca, 2018) to describe structural patterns. For example, ploting a bivariate map with the mean out degree fraction (meanODF) paired with its standard deviation (stdODF) allows to explore different situations regarding the openness of communities and the cooperation between groups of Ululers (Dao, Bothorel & Lenca, 2017). However, Hub dominance and Transitivity are particularly relevant when one considers internal patterns of organization like cooperation, because their combination leads to well-known patterns depicted on Figure 3:

- hub dominance: Internal edges of a community can be distributed in various ways around its nodes, either concentrating around a few highly centralized nodes, or uniformly distributed over the nodes. The Hub dominance metric identifies the level of centralized organization around well-connected nodes. The higher this metric of a community, the more likely it has a hub-like structure. Hub dominance can be considered as a normalized version of degree centrality. High Hub dominance leads to the well-known star-based patterns as depicted on Figure 3;
- transitivity: Very similar to the clustering coefficient (Dao, Bothorel & Lenca, 2018), Transitivity reflects the probability that adjacent vertices of a vertex are connected. This metric is usually employed to evaluate modular structures (grids) or clique dominance in networks (Fig. 3). For example, high Transitivity coupled with similar spheres of interests (or other attributes) among individuals often indicates the existence of social homophily, especially in online groups, also known as the proverb "birds of a feather flock together".

Figure 2 plots the communities in the Hub dominance vs Transitivity space. One can see that large communities are concentrated in the same area. The methods responsible of those large communities (e.g. Louvain) indeed produce a priori very few different structural patterns. With low Hub dominance and low Transitivity, most communities could be considered as "string-based" structures (Fig. 3a). In order to detect whether or not cooperation exists within large communities, we should have to zoom-in to extract dense sub-zones, i.e. apply again a detection of community to each community, and then project the new smaller communities in our bivariate map.

Conversely, Edge Betweenness, Walktrap and Fastgreedy which produce small and medium-sized communities seem to generate various directly observable types of organizations. Indeed the points are distributed in 3 of the 4 areas of our bivariate map in Figure 2. We especially find a lot of groups in the upper part of the map. This means that their members are organized around hubs, but in 2 different manners. When the Transitity is high, we find some clique-based organizations (Fig. 3d), where Ululers are (nearly) all contributing with each other to common projects. When the Transitivity is low, the organizations are mimicking star-based structures, with a very high centralization (Fig. 3c). Ululers in these groups are less involved in horizontal cooperation, but seem to follow influencers (Ululers with large degrees) who concentrate common projects with a lot of poorly connected contributors.

With these new insights, Edge Betweenness and Walktrap seem to be very good candidates: (i) they belong to the consensual group of methods previously shown, producing a robust partition; and (ii) they offer various internal organizational forms. Louvain, without offering diversified topological structures, demonstrates however interesting properties of high Hub dominance in its large communities. For this reason, we propose to keep Louvain to further explore the typology of its communities.

Fig. 2 – Structural description of communities in terms of Hub dominance and Transitivity (consensual methods)

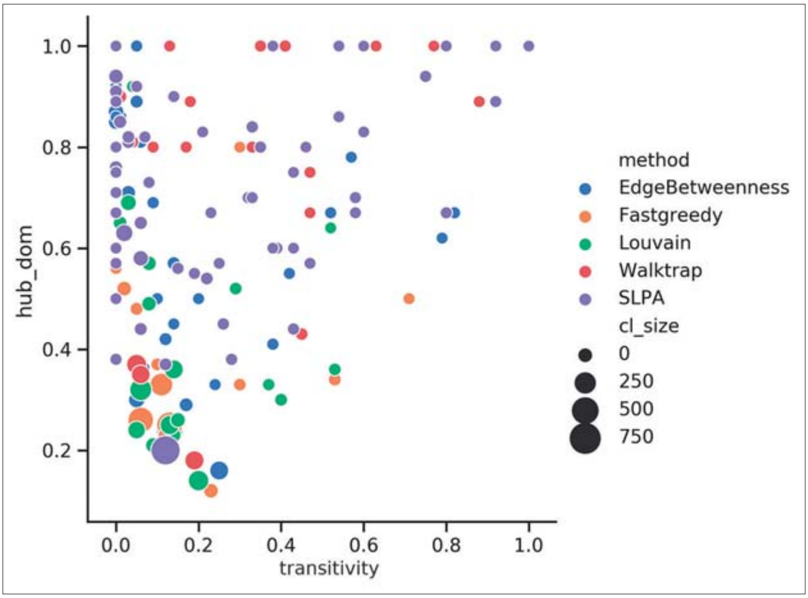
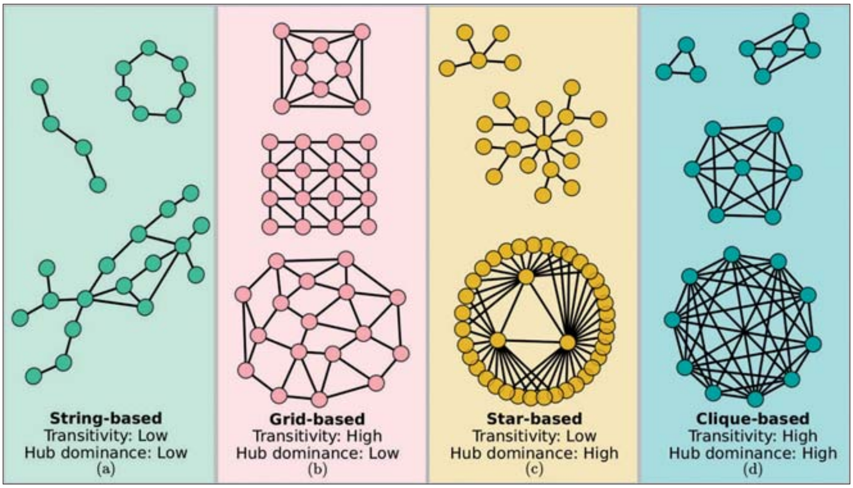


Fig. 3 – A categorization of internal community structure according to two topological property dimensions: Hub dominance and Transitivity [6]. Four representative topological communities resulting from their corresponding scores of Hub dominance and Transitivity: (a) String-based, (b) Grid-based, (c) Star-based, (d) Clique-based



To summarize, Table 2 describes the obtained communities using several conventional network metrics. First, the three partitions do not appear very distinguishable with regards to these traditional indicators. Interestingly, it seems that the size of the groups does not really influence their average degree. Moreover, the average closeness centrality and the average density also remain stable regardless of the algorithm used. Second, this confirms the need for more qualitative measures, such as the bivariate map that we proposed here above, to cover the specificities of different community detection methods and to be able to choose among them.

Tab. 2 – Conventional topological metrics to describe communities: number of obtained communities (in brackets) and average statistics calculated on the whole set of generated communities

Method (Communities count)	*Member count (mean)*	*Degree (mean)*	*Clustering coefficient (mean)*	*Betweenness (mean)*	*Closeness (mean)*
Louvain (22)	90.48	3.63	0.26	2,871.45	0.24
Edge Bet. (72)	28.90	2.96	0.20	2,234.84	0.23
Walktrap (167)	12.39	2.29	0.15	1,695.03	0.22

4. Introduction of case-specific data and typology of communities

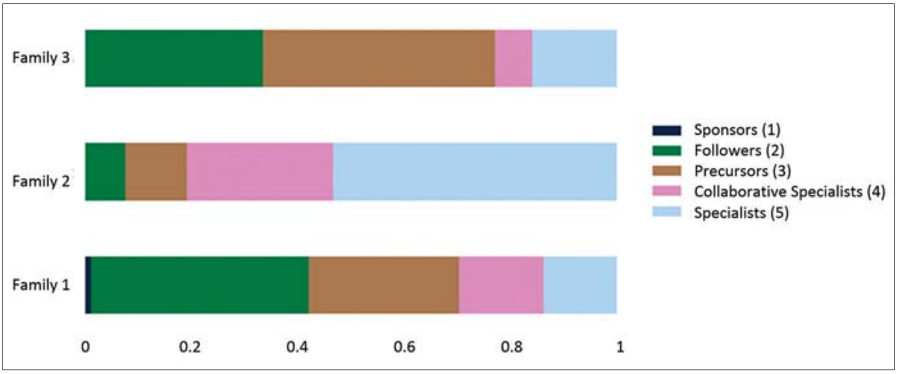
At this step, in accordance with the objective alignment approach (Smith *et al.*, 2020), we introduce additional information on the contributors' profiles, specific to the current study, which was presented in the Section 2. This helps us to compare classifications of communities, generated by three partitions, and to choose one relevant community detection method. Depending on particular context and available data, other socio-economic indicators may be mobilized for the communities' classification.

4.1. Final selection of the method

To classify the communities of Louvain, Edge Betweenness and Walktrap methods, regarding the distribution of contributors' profiles, the following clustering methods were used: (i) a principal component analysis (2 dimensions) followed by an agglomerated hierarchical clustering (Euclidean distance, Ward method of variance minimization), (ii) the K-means method and (iii) a decision tree.

The decision tree produces clusters of communities (families in Figure 4) almost identical to the K-means method (completeness = 0.964; Adjusted Rand Index: ARI = 0.854). Proximity to clusters produced by the component analysis is not obvious when considering these measures (completeness = 0.513; ARI = 0.398), but nevertheless, the distribution of the types of contributors remains relatively close.

Fig. 4 – Typology of communities (Edge Betweenness, Decision Tree)



We observe similar distributions of contributors' profiles across the different community families when we compare the 3 community detection methods and their clusters of communities. The same variety of community forms can be seen:

– Family 1 (with the Sponsors): very large balanced communities composed of all the profiles. They have the particularity of having attracted the Sponsors and a lot of Followers;
– Family 2 (Specialists): communities very clearly dominated by Specialists, whether collaborative or not;
– Family 3 (Followers, Precursors): small and even micro communities rather dominated by Precursors and to a lesser extent Followers.

Therefore, regardless the clustering method and the community detection method, the three afore-mentioned families of communities are clearly detected. The only distinguishing feature is the representation of each community family depending on the community detection method. Edge Betweenness proposes the most balanced clusters (Table 3), while Louvain produces mainly Family 1 and 2 items, and not surprisingly, most of the 167 (small) communities produced by Walktrap are from Family 3.

On this basis we choose the Edge Betweenness partition, composed of 72 communities of various sizes, exhibiting substantial examples of each family (see the exact distribution of its Ululers' profiles in Figure 4).

Tab. 3 – Sizes of communities and their contributions by Family (Edge Betweenness, Decision Tree)

Family	*Community (count)*	*Member count per community (mean)*	*Member contribution (mean in €)*	*Project goal (mean in €)*	*Project count per member (mean)*	*Project count per community (mean)*
1	8	161.5	41.4	8,050	14.4	469.5
2	28	19.1	47.7	8,501	13.3	124.8
3	36	7.0	41.3	9,589	13.7	42.5

4.2. Community structures and projects success

Once we know the community detection method, it is possible to describe the communities and their success rates in fundraising campaigns.

Table 4 uses relational (clustering coefficient, the average of the betweenness and closeness centralities of the members, mean degree) and socio-economic attributes (volume of funding, number of interactions via comments, thematic specialization and project success rate) in order to characterize collaboration rates and organization of communities.

Tab. 4 – Communities' characteristicss

Family	*Member count per community (mean)*	*Project count*	*Shared project count (mean)*	*Theme count per community (mean)*	*Comment count per project (mean)*	*Community Clustering Coefficient (mean)*
1	161.5	1,642	4.3	13.4	73.2	0.25
2	19.1	781	3.7	4.1	137.3	0.32
3	7.0	547	3.4	5.8	176.0	0.15

The three families of communities differ crucially in terms of collaboration. On crowdfunding platforms, it can take different forms: sharing projects within a community (weight of links in the Ulule social network), cohesion of members of a community around the same projects (clustering coefficient), or communication via a feedback system (comments). As presented in the Table 4, each family of communities in the case of the Ulule platform favors one of these forms.

The combination of two collaborative aspects – the weight of links in the graph (average number of shared projects) and the clustering coefficient – can inform us about community organization. For example, in Family 1, the fact that the average number of shared projects is very high but the members are poorly connected to each other (low clustering coefficient) highlights a highly centralized organization around a few central actors. We probably find here the communities with low Transitivity and high Hub dominance (probably star-based communities as seen previously). The Sponsors may play the structuring roles of these communities.

The communities in Family 2 are made up of members who share the same interests. These thematic groups, which are strongly connected and supportive, collectively take their decisions on project funding and contribute on average more than other communities.

Finally, in Family 3, communities are also structured around certain themes but without clear specialization. Members of these communities communicate a lot through the feedback system (comments).

4.3. Comparative analysis of success rates of families of communities

Since the thematic categories are not represented in a balanced way in the three Families of Communities, we cannot directly compare the respective effectiveness of community forms for the different themes. Nevertheless, if we look in more detail at the communities that have a 100 percent success rate for all our indicators (Table 5), there are 13 communities belonging to Families 2 and 3 (respectively 9 and 4 communities). These two Families have very close thematic choices.

Tab. 5 – Communities with high success rates vs. communities with lower success rates. The indicators are averaged by community. Sub-table "All" shows the indicators for all 72 communities

	Shared project count (mean)	*Comment count per project (mean)*	*Project goal (mean)*	*Member count (mean)*	*Theme count (mean)*	*Degree (mean)*	*Betweenness (mean)*	*Clustering coefficient (mean)*
				Success Rate 100% (13 communities)				
mean	11	517	8,826	6	4	2.84	2,126	0.39
std	6	1,120	6,132	3	3	1.06	1,356	0.35
min	3	2	2,630	3	1	1.67	777	0
25%	6	21	4,408	3	1	2.00	1,419	0
50%	9	60	7,946	4	3	2,50	1,559	0.43
75%	15	123	10,807	6	7	3.25	2,078	0.67
max	22	3 840	25,019	13	9	4.85	5,869	0.88
				Success Rate < 85% (9 communities)				
mean	11	234	9,126	4	4	2.06	1,796	0.09
std	10	534	6,271	2	3	0.85	721	0.17
min	5	15	2,202	2	2	1.50	693	0
25%	6	29	6,305	3	3	1.67	1,039	0
50%	6	35	7,006	3	3	1.80	2,078	0
75%	12	88	11,189	5	3	2,00	2,078	0.06
max	35	1,653	23,916	8	11	4.25	2,970	0.44
				All (72 communities)				
mean	58	393	10,832	29	6	2.96	2,235	0.20
std	122	686	5,799	78	4	1.18	1,092	0.22
min	3	2	2,202	2	1	1.50	693	0
25%	8	33	6,894	3	3	2.00	1,558	0
50%	17	78	9,053	8	5	2.93	2,077	0.15
75%	36	381	14,915	13	8	3.63	2,728	0.30
max	789	3,840	27,555	579	15	6.46	5,869	0.88

These best communities are small in size, quite specialized and finance projects on average on 4 different themes. The projects are very different in terms of the objectives to be achieved (all types of amounts are represented), but less than 25% of the projects have a larger than average scope, so we have rather modest projects.

Surprisingly, the 9 least successful communities in the graph also belong to Families 2 and 3. What can be noticed in the Table 6 is that the most and the least successful communities share quite similar characteristics (size, thematic specialization). What clearly differentiates the best performing communities is the clustering coefficient and the number of comments, which are significantly higher. In other words, the thematic specialization that characterizes Family 2 and partially Family 3, does not in itself guarantee success of crowdfunding projects. To reach significant economic performance of fundraising campaigns, they must be coupled with strong social involvement and cohesion of its members.

With regard to Family 1, characterized by a high centralization and a high thematic openness, we notice a relatively high success rate. Thematic diversity also attracts participants in communities.

5. Conclusion

Community detection makes it possible to identify very diverse groups in a social network. This paper demonstrates a methodology to choose one relevant community detection algorithm, among 11 well-known ones, providing fruitful insights into the cooperation forms not directly observable on a crowdfunding platform.

The choice of a particular community detection method is not an easy or a neutral choice. As demonstrated in the paper, depending of partition methods, practitioners obtain a range of various community types, which will drastically change the final results of their analysis. This paper substantiates that an accurate way to choose one suitable method is a complex task. Especially in the context of exploratory studies it necessitates the combination of a range of techniques, e.g. in our case, partitions' similarities, qualitative criteria and structural indicators (string- based, star-based or clique-based organizations of communities).

In line with N. Smith *et al.* (2020) this study substantiates that the choice of a method is determined by the research context and problematics. Additional techniques, specific data and indicators allow to narrow down the scope of available options in the methods choice. Their alignment with the practitioners' research question plays a crucial role for the final choice of a particular method. In the framework of the case study presented in this paper, the choice of the Edge Betweenness method results from the analysis of socio-economic characteristics and the exploration of the distribution of Ululers' profiles. This way, we have identified 3 robust families of platform's communities and their distinctive features, i.e. organization, number of participants, collaboration intensity, thematic specialization, and performance in the fundraising campaigns. Depending on the context and available data, different socio-economic indicators may be mobilized to obtain the communities' classification and a range of further business-oriented questions may be addressed: e.g. precise distribution of string-based, star-based or clique-based forms in communities' families, life circle and evolution dynamics of the communities and many others.

References

Agrawal A., Catalini C., Goldfarb A. (2010), "Entrepreneurial finance and the flat world hypothesis: Evidence from crowdfunding entrepreneurs in the arts", *Working Papers 10-08*, NET Institute.

Blondel V.D., Guillaume J.L., Lambiotte R., Lefebvre E. (2008), "Fast unfolding of communities in large networks", J*ournal of Statistical Mechanics: Theory and Experiment*, 10, 10008.

Chakraborty T., Dalmia A., Mukherjee A., Ganguly N. (2017), "Metrics for community analysis: A survey", *ACM Comput. Surv.*, 50 (4), pp. 1-37.

Clauset A., Newman M.E.J., Moore C. (2004), "Finding community structure in very large networks", *Physical Review*, 70 (6).

Dao V.L., Bothorel C., Lenca P. (2017), "Community structures evaluation in complex networks: A descriptive approach", in E. Shmueli, B. Barzel, R. Puzis (*eds.*), *3rd International Winter School and Conference on Network Science, NetSci-X 2017*, Springer International Publishing, Cham, pp. 11-19.

Dao V.L., Bothorel C., Lenca P. (2018), *An empirical characterization of community structures in complex networks using a bivariate map of quality metrics*, CoRR abs/1806.01386.

Dao V.L., Bothorel C., Lenca P. (2019), "Estimating the similarity of community detection methods based on cluster size distribution", in L.M. Aiello, C. Cherifi, H. Cherifi, R. Lambiotte, P. Lió, L.M. Rocha (*eds.*), *Complex Networks and Their Applications VII*, Springer International Publishing, Cham, pp. 183-194.

Dao V.L., Bothorel C., Lenca P. (2020), "Community structure: A comparative evaluation of community detection methods", *Network Science*, 8 (1), pp. 1-41.

Fortunato S., Hric D. (2016), "Community detection in networks: A user guide", *Physics Reports*, 659, pp. 1-44.

Girvan M., Newman M.E.J. (2002), "Community structure in social and biological networks", *Proceedings of the National Academy of Sciences*, 99 (12), pp. 7821-7826.

Inbar Y., Barzilay O. (2014), *Community impact on crowdfunding performance*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2524910.

Jebabli M., Cherifi H., Cherifi C., Hamouda A. (2018), "Community detection algorithm evaluation with ground-truth data", *Physica A: Statistical Mechanics and Its Applications*, 492, pp. 651-706.

Karrer B., Newman M.E.J. (2011), "Stochastic blockmodels and community structure in networks", *Physical Review*, E 83 (1).

Kuppuswamy V., Bayus B.L. (2018), "Crowdfunding creative ideas: The dynamics of project backers", in *The Economics of Crowdfunding*, Springer International Publishing, Cham, pp. 151-182.

Lyubareva I., Brisson L., Bothorel C., Billot R. (2020), "Une plateforme de crowdfunding et son réseau social: l'exemple ulule", *Revue Francaise de Gestion*, 1 (286), pp. 135-151.

Meo P.D., Ferrara E., Fiumara G., Provetti A. (2014), "Mixing local and global information for community detection in large networks", *Journal of Computer and System Sciences*, 80 (1), pp. 72-87.

Newman M.E.J. (2006), "Modularity and community structure in networks", *Proceedings of the National Academy of Sciences*, 103 (23), pp. 8577-8582.

Newman M.E.J., Girvan M. (2004), "Finding and evaluating community structure in networks", *Physical Review*, E 69 (2).

Pons P., Latapy M. (2005), *Computing communities in large networks using random walks (long version)*, ArXiv Physics e-prints.

Raghavan U.N., Albert R., Kumara S. (2007), "Near linear time algorithm to detect community structures in large-scale networks", *Physical Review*, E 76 (3).

Reichardt J., Bornholdt S. (2006), "Statistical mechanics of community detection", *Physical Review*, E 74 (1).

Rosvall M., Bergstrom C.T. (2008), "Maps of random walks on complex networks reveal community structure", *Proceedings of the National Academy of Sciences*, 105 (4), pp. 1118-1123.

Schaub M.T., Delvenne J.C., Rosvall M., Lambiotte R. (2017), "The many facets of community detection in complex networks", *Applied Network Science*, 2 (1).

Smith N.R., Zivich P.N., Frerichs L.M., Moody J., Aiello A.E. (2020), "A Guide for Choosing Community Detection Algorithms in Social Network Studies: The Question Alignment Approach", *American Journal of Preventive Medicine*, 59, 4, pp. 597-605.

Xie J., Szymanski B.K. (2011), "Community detection using a neighborhood strength driven label propagation algorithm", *Proceedings of the 2011 IEEE Network Science Workshop, NSW*, 11, pp. 188-195.

Yang Z., Algesheimer R., Tessone C.J. (2016), "A comparative analysis of community detection algorithms on artificial networks", *Sci Rep.*, 6, 30750.

Zheng H., Li D., Wu J., Xu Y (2014), "The role of multidimensional social capital in crowdfunding: A comparative study in China and us", *Information & Management*, 51 (4), pp. 488-496.

2. Examining the association between behavioral trajectory similarity clusters in social networks[*]

by Brandon Sepulvado[**], Omar Lizardo[***], Mike Wood[****], Cheng Wang[*****], David Hachen[******]

1. Introduction

The increasing availability of fine-grained temporal data on human interaction in social networks (Bahulkar *et al.*, 2017; Kim & Anderson, 2012; Liu *et al.*, 2018; Miritello, 2013; Sekara *et al.*, 2016), augmented with equally fine-grained information on attitudes, habits, and practices collected via unobtrusive means (Purta *et al.*, 2016; Purta & Striegel, 2019), has opened up new opportunities to study the link between social networks and human behavior (Lazer *et al.*, 2009). For social networks, rather than thinking of pairwise interactions in the model of a static graph, there is now an emphasis on temporal dynamics and developing temporal versions of quantitative network structure (Dickison *et al.*, 2016; Holme & Saramäki, 2012). In the study of human behavior, there is a renewed emphasis on methods attentive to dynamic change, such as event-history and time-series analysis (Epskamp, 2020).

[*] Research reported in this paper was supported in full by the National Heart, Lung and Blood Institute of the National Institutes of Health under award number R01HL117757 ($2,913,061). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[**] Department of Methodology and Quantitative Social Sciences, NORC at the University of Chicago, 4350 East-West Highway, 8th Floor, Bethesda, MD 20814, USA, sepulvado-brandon@norc.org.

[***] Department of Sociology, University of California, Los Angeles, 375 Portola Plaza, 264 Haines Hall, Los Angeles, CA 90095, USA, olizardo@soc.ucla.edu.

[****] Department of Sociology, Brigham Young University, 2008 JFSB, Provo, UT 84602, USA, michaelwood@byu.edu.

[*****] Department of Sociology, 2257 Faculty Administration Building, Wayne State University, Detroit, MI 48202, USA, chengwang@wayne.edu.

[******] Department of Sociology, University of Notre Dame, 4076 Jenkins Nanovic Hall, Notre Dame, IN 46556, USA, David.S.Hachen.1@nd.edu.

Our recent work has attempted to join these lines of work by taking the classic social network concept of dyadic behavioral similarity and looking at it from a temporal perspective that emphasizes dynamic trajectories (Sepulvado *et al.*, forthcoming). We depart from most prior work on similarity, which has looked at it statically with the concept of homophily. There has been a great deal of research attempting to both predict various types of homophily and assess its impact on network outcomes (Aral *et al.*, 2009; Kossinets & Watts, 2007; Ingram & Morris, 2007; Yang *et al.*, 2019; Lewis & Kaufman, 2018; Leszczensky & Pink, 2019). However, there has been very little work looking at how changes in similarity predict tie formation, and how tie formation results in changes in the similarity between two persons in their behaviors. In this respect, the "dynamic" turn toward the study of dyads in general and homophily in particular advocated by Rivera *et al.* (2010) about ten years ago has yet come to pass (but see Bahulkar *et al.*, 2017; Dokuka *et al.*, 2016; Schaefer & Kreager, 2020; Wang *et al.*, 2020).

2. Background

In an earlier paper (Sepulvado *et al.*, forthcoming), we developed the idea of dyadic similarity trajectories and proposed methods for assigning dyads into discrete, substantively meaningful clusters that help predict dyadic connectivity and dyadic matching (homophily) on key sociodemographic traits. This paper uses those methods to explore the predictive linkage between two similarity trajectories, one based on physical activity and the other on social activity.

The larger goal of this line of work is to model the association between various similarity trajectories in order to improve our understanding of how social ties in social networks form, evolve, and decay. In this paper, we focus on dyadic similarity trajectories based on similarity in the two vertices' network position as measured by their *temporal outdegree*, the number of people that they communicate with per unit of time and an indicator of how socially active a person is in temporal social networks (Holme & Saramäki, 2012).

Of particular interest is the correspondence between trajectories based on similarity in network position and trajectories based on similarity in behavior; in our case, physical activity is measured by daily step counts. We use these two measures of similarity, how active people are physically and socially, to ascertain the extent to which similarity in outdegree is associated with both the probability of a tie and similarity in physical activity. Insights gained by analyzing the association between similarity in social and physical activity will enhance our understanding of how behavioral similarity trajectories can help predict tie formation, evolution, and persistence.

We use data from the longitudinal *NetHealth* Study, which collected network, behavioral, and attitudinal data from smartphones, activity trackers, and surveys. From this rich-attribute, fine-grained longitudinal social network and behavior dataset, we generate dyadic communication similarity trajectories using daily counts of the number of outgoing (person-initiated) communication events obtained from smartphone logs. We compute daily similarity scores for each dyad, which we use to create a final "second-order" dyadic communication similarity trajectory. We then use clustering techniques sensitive to the temporal (i.e., time-series) structure of the data to generate dyadic clusters, showing that the clusters help predict dyadic connectivity. Finally, we look at the statistical association between dyadic clusters based on communication trajectories with clusters based on physical activity trajectories, and we ask whether this association is moderated by dyadic matching in sociodemographic traits.

2.1. Key Concepts

We begin by defining some key terms. This includes the network notion of a *dyad*, the idea of a *behavioral trajectory*, the concept of *behavioral trajectory similarity clusters* defined over dyads, and the *association* between distinct behavioral trajectory similarity clusters when considering two or more behaviors, which may include endogenous communicative activity in social networks. In the discussion, we elaborate on how our approach can be extended to other behavioral and non-behavioral temporal patterns of change and higher-order network motifs beyond dyads.

2.2. Dyads

We begin with the classic social network concept of the *dyad* (Wasserman & Faust, 1994, p. 505-ff). In social networks represented as a graph G with a set of undirected links *E*, a set of vertices *V*, and associated symmetric adjacency matrix **A**, a dyad is defined as all unordered pair of actors in the system (in directed networks, the pairs are ordered); each cell ($\mathbf{a}_{ij}$) in either the upper or lower triangle of the symmetric adjacency matrix ($\mathbf{a}_{ij} = \mathbf{a}_{ij,\ \forall i,\ j \in V}$) refers to a dyad in the network. In the simplest case of an undirected network, there are two mutually exclusive types of dyads: *connected* (in which case $\mathbf{a}_{ij} = 1$)

and *null* (in which case $\mathbf{a}_{ij} = 0$). Connected dyads are joined by an edge (link or tie) in the social network, whereas null dyads do not share an edge.

In the temporal network case, dyads must be indexed at each timepoint because actors may join or leave the system and because dyads may transition from one state (e.g., similar) to another (e.g., dissimilar) or from being connected to null (and back). Accordingly, the question of whether two dyads are similar on a given trait must be separated from whether they have a direct link in the temporal network. More formally, at any point in time (t), an edge between two actors, i and j, may or may not exist, which can be written, using matrix notation, as $\mathbf{a}_{ijt} = 1$ if the edge exists at time t, $\mathbf{a}_{ijt}$ = or 0 if the edge does not exist at time t. A dyad can be similar or dissimilar on a given time-varying trait and be either connected or disconnected in the temporal network. Whether connected dyads are more behaviorally similar is an empirical question.

2.3. Actor-Level Behavioral Trajectories

Each actor in a temporal social network is observed at multiple points in time. This means that information obtained on behaviors, traits, and habits of the actor can also dynamically change over the observation period (Christakis & Fowler, 2007; Fowler & Christakis, 2008; Lazer *et al.*, 2009; Lewis *et al.*, 2008). Consider the value of a given trait or behavior s, observed for actor i at time t. The time series of values $s_i = \{s_{it(1)}, s_{it(2)}, s_{it(3)} \ldots s_{it(m)}\}$ across all time points m defines a *behavioral trajectory* for that actor on that trait. For instance, if the trait is something like physical activity, such a trajectory might indicate that an actor's physical activity level might increase, decrease, or vacillate between the two. In essence, behavioral trajectories can take multiple functional forms with respect to time (Sepulvado *et al.*, forthcoming).

2.4. Dyad-Level Behavioral Trajectory Similarity

Since each actor in the network has a behavioral trajectory, each dyad can be more or less similar on that trait at each point in time. More formally, for each pair of actors in a temporal network i and j, with behavioral trajectories on trait s, denoted by s_i and s_j, we can define the dyadic behavioral trajectory similarity at each point in time s_{ijt} as the absolute value of the difference between the value of the trait for each actor at that time, or $|s_{it} - s_{jt}|$. Note that just like for each behavioral trajectory, the dyadic similarity trajectory defines a time-series over each observed time point for each dyad in the network, $s_{ij} = \{s_{ijt(1)}, s_{ij(t2)}, s_{ijt(3)} \ldots s_{ijt(m)}\}$.

This time series may itself display substantively meaningful behavior. For instance, a dyad may become more similar or dissimilar on the observed trait or display fluctuating similarity and dissimilarity patterns over time. Dyadic behavioral trajectory similarities and dissimilarities may also be affected by a variety of dyad-level factors, inclusive of dyadic similarity on other (both time-varying and -invariant) traits, such as whether the two actors in the dyad have similar value on sociodemographic attributes (e.g., race, religion) or environmental factors (e.g., propinquity) relevant to connectivity in social networks (Feld, 1982; McPherson *et al.*, 2001; Rivera *et al.*, 2010; Sepulvado *et al.*, forthcoming).

We can link the dynamics of dyadic behavioral trajectory similarity with other dynamics in a social network, most importantly the dynamic *transition* of dyads from null to connected (or vice versa) and, in directed networks, dynamic transitions of dyads from null, to asymmetric, to mutual (Wasserman & Faust, 1994, p. 505-ff). This can advance core issues in social network analysis, such as whether similarities precede connectivity (transition of dyad from a null to a connected state), whether connectivity is a causal input into increasing similarity, and whether increasing dissimilarity leads to disconnectivity (transition of a dyad from a connected to null state) or vice versa (Bahulkar *et al.*, 2017; Lewis & Kaufman, 2018; Noel & Nyhan, 2011; Schaefer & Kraeger, 2020).

2.5. Dyad-Level Behavioral Trajectory Similarity Clusters

The total possible number of dyads in a network increases super-linearly in the number of actors. As is well-known, for an undirected network represented by a graph of order *n*, there are $\frac{1}{2}n(n-1)$ possible dyads (Wasserman & Faust, 1994, p. 515). This means that the number of dyadic similarity trajectories to be considered will similarly increase. When examining dyadic behavioral similarity trajectories in temporal networks, it is desirable to look for a way to cluster dyads into a smaller set of classes. The idea is to assign dyads to the same class when they have similar (e.g., substantively the same except for measurement error and small statistical deviations) behavioral similarity trajectories (e.g., all the dyads in the same class become more similar over time). Such a possible set of dyadic classes or clusters in social networks have been called *behavioral trajectory similarity clusters* (Sepulvado *et al.*, forthcoming).

Although previous studies have examined temporal changes in tie characteristics (Krackhardt & Hancock, 2006; Martin & Yeung, 2006; Morgan *et al.*, 1997; Schaefer & Kraeger, 2020) or the emergence of collective similarity due to the diffusion of contagions (Centola, 2010; Christakis & Fowler, 2007, 2008, 2013; Fowler & Christakis, 2008b, 2008a), the concept of behavioral similarity trajectory clusters solves two fundamental problems (Sepulvado *et al.*, forthcoming). First, many studies researching similarity dynamics in social networks condition on a dyad already being connected (Fowler & Christakis, 2008b, 2008a), leading to biased conclusions about how ties impact dyadic similarity (e.g., Cohen-Cole & Fletcher, 2008). Second, similarity on behavioral and/or attitudinal traits is often considered at only a single time point (e.g., Aral *et al.*, 2009; Feiler & Kleinbaum, 2015; Lewis *et al.*, 2012), yet this stasis is not a valid assumption for many individual characteristics. For example, political attitudes can change, and behavioral patterns evolve with life events (Bidart & Lavenu, 2005), and as seasons change. Beyond addressing common issues with network analysis studies, such trajectory clusters enable theoretical innovation.

In this paper, we extend our previous work in two significant ways. First, we extend the notion of behavioral trajectory similarity from exogenous (e.g., physical activity) to *endogenous* social network traits. Namely, we use the *communicative activity* level to define a behavioral temporal degree trajectory for each actor in the network (Miritello, 2013; Miritello *et al.*, 2013; Raeder *et al.*, 2011). In a temporal network, each actor i may communicate (send ties) to a set of actors k at time t. The quantity $k(i)_t$ defines the level of communication of that person at that time, a measure of temporal network centrality referred to as *temporal outdegree* (Kim & Anderson, 2012). This measure of centrality is an indicator of a person's sociability and varies both between persons (in time-averaged slices) and within persons over time.

Absolute temporal degree *differences* between the members of the dyad in the network ($s_{ijt} = |k(i)_t - k(j)_t|$) thus define a set of behavioral trajectory similarity series based upon communication for each dyad. These, in turn, can be used to assign each dyad to a behavioral trajectory similarity cluster, as defined earlier. The second way in which we extend our previous work is by looking at multiple behavioral trajectory similarities in the same analysis. To that end, we construct behavioral trajectory similarity clusters based on an exogenous trait (daily step counts) and an endogenous trait (temporal degree). We then look at whether these two dyadic class assignments are statistically dependent upon one another using methods to detect a statistical association in categorical (polytomous) variables (Powers & Xie, 2008). In this way, we can examine the linkage between multiple dyadic behavioral trajectory similarity clusters.

3. Data

We use data from a study called NetHealth that followed a cohort of 625 undergraduates at the University of Notre Dame (Purta *et al.*, 2016). To examine the relationship between health behaviors, communication activities, social networks, and other actor-level traits, students were equipped with smartphones and activity trackers worn on the wrist (i.e., Fitbits) and surveyed before matriculation and every semester after that. Although the entire study period was from Fall 2015 to Spring 2019, only the Fall 2015 data are included in these analyses. We exclude data before 1 September 2015 because students were still picking up their Fitbit devices.

Communication data come from smartphones, specifically "metadata" from calls, text messages, and WhatsApp messages. The metadata includes timestamped information of communication events and the numbers of the caller/sender and receiver. We use the data to construct a temporally aggregated social network among study participants using a threshold: treating an edge as existing if there was any communication attempt in the period between any two pairs of actors i and j. For physical activity data, we use daily steps as measured by the Fitbits.

The communication and physical activity data are used to create distinct trajectories: first at the actor level and then for dyads. Each participant has a number of initiated communication events for each day in the 109-day Fall 2015 period. This daily communication activity is directed. We use this information to compute each person's temporal (daily) degree in the network, defined as the number of other actors in the network contacted at time t. Similarly, each person has a daily total number of steps for each day in the period. Actor-level behavioral trajectory time series are constructed from both daily communications and steps. To derive dyadic similarity trajectories, we compute the absolute daily difference in degree and steps for each dyad.

There are 195,000 possible dyads in the NetHealth Study network, but not all of these are included in analyses. If at least one of the actors in a dyad had missing data for a given day, we treated the dyad's steps as missing for that day. Study participants could have no missing data on communication activity because days without communication were coded as zero. We excluded dyads that were missing over 25% of their daily step differences, and, for those dyads with up to 25% missing data, conducted linear interpolation. After filtering out missing data, 32,872 dyads remain.

Participants completed surveys collecting demographic information used in the analysis. We use this information to create three binary variables indicating whether both actors in a dyad have the same responses for race/ethnicity, gender, and religious identity. The final sample of 282 participants includes 156 male, 126 female students, 192 students identifying as white, 15 students identifying as Black, 25 identifying as Asian, 36 identifying as Latino, one student who did not identify with any of the racial identities, and 13 who were born outside the U.S. The sample includes 211 Catholics, 36 Protestants, 10 students identifying with "Other Religion" and 25 identifying as not religious. 16,563 dyads identify as the same gender, 16,602 identify with the same race/ethnicity, and 19,359 identify with the same religious identity.

4. Methods

To identify dyadic *classes* of (dis)similarity trajectories, we turn to unsupervised learning. Yet, given that we start with trajectories (i.e., time series) of 109 daily differences in steps and communication activity for all dyads, most traditional clustering methods are unsuitable. We thus rely upon time series clustering methods. There are three general approaches within this class of clustering methods: clustering on specific time points, clustering on subsequences of observations, or clustering the entire time-series (Aghabozorgi *et al.*, 2015, pp. 18-19). We choose the third strategy because we are interested in the evolution of dyadic similarity across the full Fall 2015 semester.

Rather than clustering using model-based methods (e.g., clustering on output from ARIMA models) or feature-based methods (e.g., clustering lower-dimensional representations of time series), we use a shape-based method that clusters multiple time series based upon the similarity of their shapes because previous evaluations have demonstrated that shape-based methods exhibit superior performance (Aghabozorgi *et al.*, 2015, pp. 18-19; Liao, 2005; Paparrizos & Gravano, 2017). Because it outperforms various alternatives for time series data (Paparrizos & Gravano, 2015, 2017), we choose the *k*-shape algorithm. *k*-shape clustering is based on the *k*-means algorithm and considers multiple known distortions that are frequently present in time-series data (Paparrizos & Gravano, 2015, p. 1859). For example, time-series' features might be scaled differently (i.e., scaling invariance), and specific subsequences of the two time-series might have a similar shape while other subsequences considerably diverge (i.e., shift invariance).

We assess the meaningfulness and validity of the dyadic trajectory similarity clusters, based on temporal degree and obtained using the method described above in two ways. First, we examine whether dyads assigned to the same temporal degree trajectory similarity class are more or less likely to transition from the null to the connected state during the observation period.

For this analysis, we specify a model predicting a dyad's probability of being connected from the assigned similarity trajectory cluster membership and compute predicted probabilities of having a tie for each cluster. We use logistic regression with Firth (1992) penalization to account for any bias produced by the "rare-event" nature of the dyadic outcome (there is a larger class imbalance, with null dyads outnumbering connected dyads by a factor of 49 to 1). Additionally, we tested a Firth-penalized model with intercept correction (not presented; Puhr *et al.*, 2017), which did not substantively change model performance. The Firth logistic regression treats whether a dyad has a tie (i.e., 1 if the two individuals communicated during the Fall 2015 semester) as the dependent variable and cluster membership as the dependent variables.

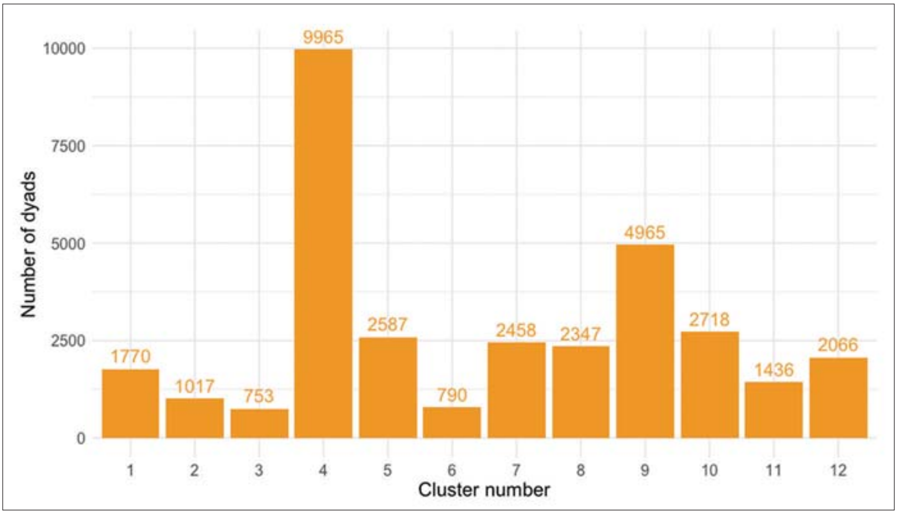
Second, and this is the crucial innovation introduced in the paper, we examine the question of whether dyadic trajectory similarity classes based on one (endogenous) trait (i.e., temporal degree) are associated with dyadic trajectory similarity classes based on another (exogenous) trait (daily steps). We use log-linear models to examine the association between cluster membership along these two dyadic trajectory similarity dimensions (Powers & Xie, 2008). Additionally, we examine the extent to which any association between the two dyadic clusters is due to dyadic matching on time-invariant sociodemographic factors, which has been called "homophily" in the social networks literature (McPherson *et al.*, 2001; Rivera *et al.*, 2010). To do this, we use log-linear models of the three-way association between dyadic temporal degree and step similarity trajectory cluster, and three types of sociodemographic homophily – gender, race, and religion. These models allow us to ascertain the extent to which the association between step and communication activity clusters is partially the result of the association between demographic similarity and both step- and communication-based dyadic similarity trajectory cluster membership.

We first estimate a Poisson model using the steps and communication clusters to predict cell count (i.e., the number of dyads falling into each combination) and identify systematic associations by examining which cells have high residuals (i.e., where cell count differs from what would be expected by chance). Next, we run a series of nested log-linear models (Powers & Xie, 2008) to examine the three-way association between demographic homophily measures (same gender, same race, or same religion), physical activity cluster membership, and temporal degree cluster membership. The baseline (independence) model fits the marginal distributions for the steps clusters, the communication clusters, and the homophily measure. The second model fits parameters for each homophily factor's association with the steps and the communication clusters. The final model fits the two-way marginal between step and temporal degree cluster membership. These models regress cell count on different specifications of the step trajectory clusters, communication trajectory clusters, and each sociodemographic variable; each model's exact specification may be found in the "Equation" column of Table 1.

5. Results

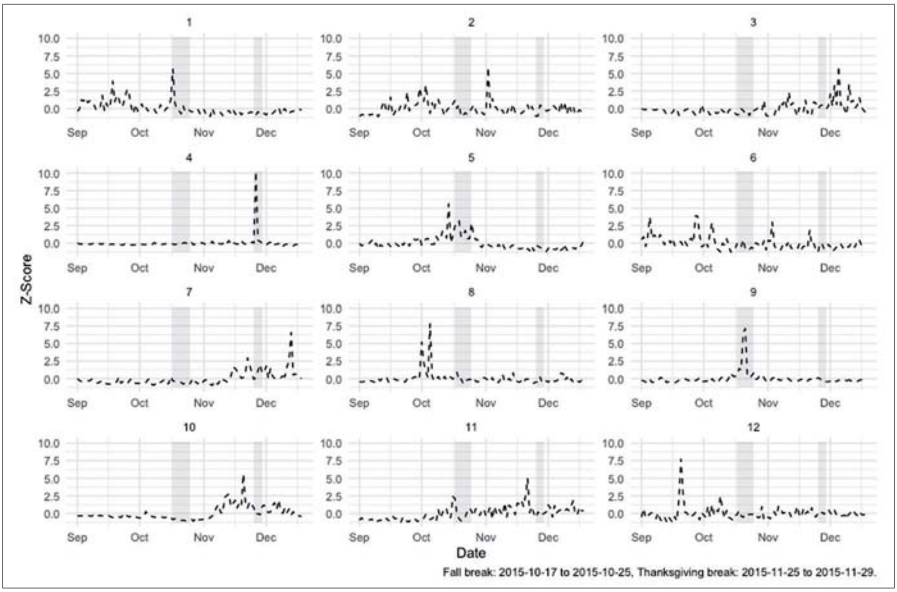
The bar plot in Figure 1 shows the sizes of the clusters derived from outgoing communication, while Figure 2 presents the typical trajectories for each dyadic class identified by the *k*-shape clustering algorithm. The line in each plot in Figure 2 represents the centroid trajectory of a cluster. The *x*-axes indicate the day (with month labels), and the *y*-axes indicate the *z*-normalized value for the absolute difference in temporal degree for all the dyads in that class on that day. Recall that each dyad trajectory is a time series that was initially the absolute difference in daily activity for the actors in a dyad.

Fig. 1 – Sizes of degree-based clusters



There are two clusters (4 and 9) with a constant standardized difference score of zero across all days except for one large peak. Each peak occurs during a holiday break period where students typically leave campus and visit home, as noted by the figures' shaded area (Fall Break and Thanksgiving, respectively). These are the two largest clusters, containing 14,930 dyads (almost 50%). The peaks indicate periods when there is a great deal more communication heterogeneity in the population, which results in dyads between people with substantially different activity levels. There are also peaks evident in other cluster profiles (e.g., 8 and 12), but the *z*-score is not flat outside those peak periods.

Fig. 2 – Cluster centroids for degree-based clusters

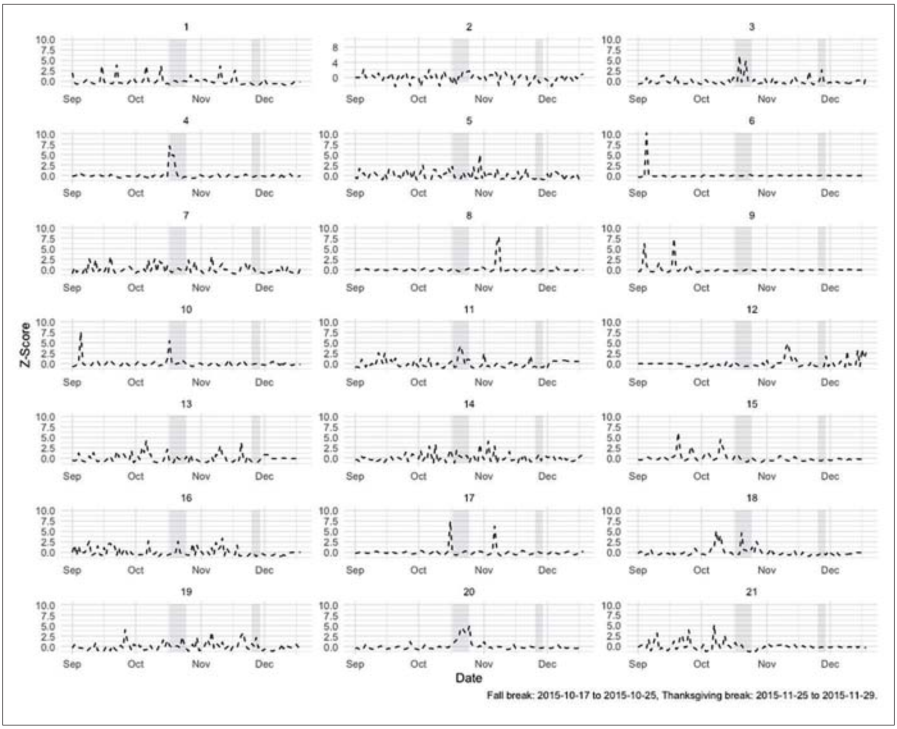


If we focus on non-peak periods, other patterns are evident. In clusters 1 and 2, there is a good deal of difference early in the period, but scores eventually stabilize (indicating convergence in temporal degree over time). For clusters 3 and 7, it is the reverse (indicating divergence). Dyads in clusters 5 and 10 exhibit an inverted-U shape pattern with a period of dissimilarity in the middle. Clusters 6, 11, and – to some extent – 12 have more of an erratic pattern with periods in which the *z*-scores are below zero, thus indicating above-average similarity between the dyads.

Figure 3 shows centroid plots for the 21 dyadic trajectory similarity clusters based upon steps. The CVIs suggested 21 steps-based similarity trajectories, and the number of dyads within each cluster is distributed much more evenly than with the communication similarity trajectories. Clusters 3, 4, 10, 16, and 18 tend to have extreme dissimilarity during the mid-semester break, though they each exhibit different combinations of similarity and similarity throughout the rest of the semester. Clusters 3 and 19 peak during the Thanksgiving break, but dyads in 3 remain otherwise much more similar than those in 19. We refer readers to Sepulvado *et al.* (forthcoming) for a more detailed description of the clusters.

Figure 4 presents the results of the logistic regression predicting the tie probability based on cluster membership. We present the predicted probabilities of a tie derived from model parameter estimates. Most predicted probabilities are around .02, which is the NetHealth social network's overall density in Fall 2015 (equivalent to the base probability of being a connected dyad). Clusters 2 and 3 have higher than expected probabilities, while dyads in clusters 10 and 11 are less likely to be connected than expected by chance.

Fig. 3 – Cluster Centroids for daily step dyadic trajectory similarity clusters



5.1. Association Between Steps and Temporal Degree Clusters

Next, we turn to the log-linear models. With residual deviance of 4664.2 on 220 degrees of freedom, the Poisson model predicting the number of dyads falling into each step and communication cluster combination has a p-value < 0.001. Figure 5 visualizes the residuals from the predicted counts under the independence model: (observed count - predicted count)/sqrt(predicted count). The largest positive residual values (> 5) are found in 6% of the 252 cells. There are fewer negative residual cells (2% with residuals < -5), and all these cases involve communication clusters which have a stable daily z-scores except for one peak (clusters 4, 5, 9, 12), suggesting it is unlikely this step cluster type is associated with a flat temporal degree difference profile.

Fig. 4 – Predicted probability of connected dyads by temporal degree cluster

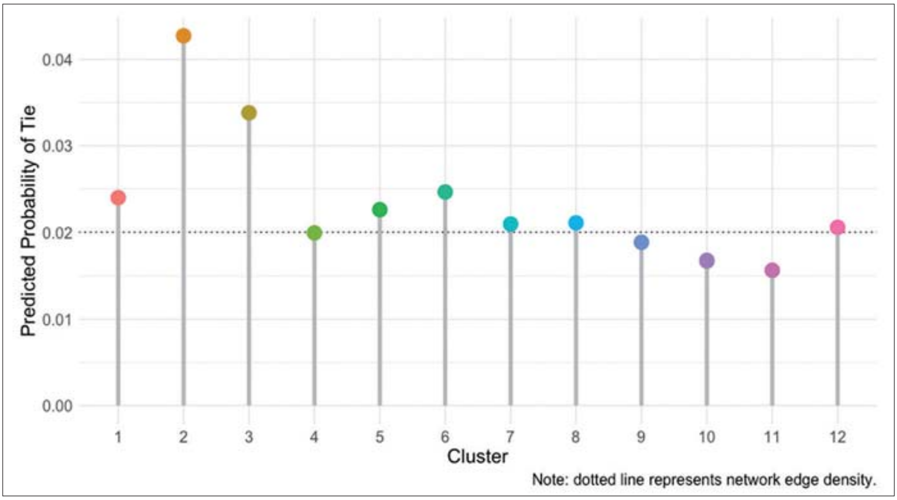
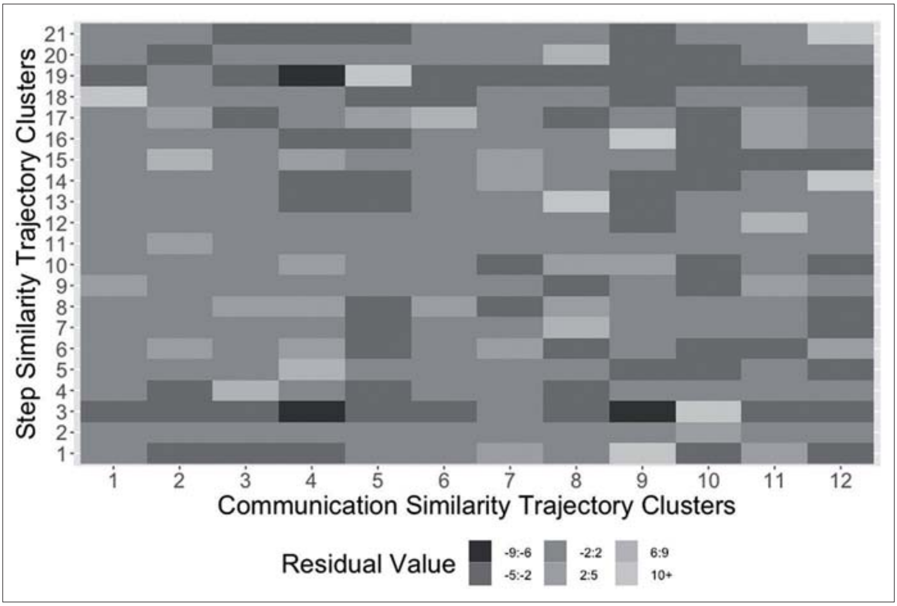


Fig. 5 – Heatmap of the association between trajectory clusters



We ran a series of nested log-linear models (Powers & Xie, 2008) to examine the three-way association between demographic homophily measures (same gender, same race, or same religion) and step and communication activity cluster membership. Table 1 shows the results for gender, race, and religious homophily. Recall that the baseline (independence) model fits the marginal distributions for the steps clusters, the communication clusters, and the homophily measure and that the second model fits parameters for each homophily factor's association with the steps and the communication clusters. As shown in the table, adding the two-way marginals that fit the association of gender, race, and religious homophilous dyads with step and communication trajectory similarity clusters statistically improves the fit of the model using conventional criteria of significance ($p < 0.01$). This indicates that dyads assigned to the same temporal degree and step trajectory similarity class are also more (or less) likely than expected by chance to match on key sociodemographic characteristics.

However, looking at the *substantive* improvement in model fit (indicated by the proportion reduction in deviance), we can see that the association between step and communication trajectory similarity clusters and race and religious homophily is much stronger than the corresponding improvement in model fit for gender (15% and 20% versus 1.6% respectively). This indicates that, while statistically discernible, the association between dyadic degree trajectory similarity cluster and gender is much weaker than for the other homophily dimensions. This is consistent with results reported in our previous work, which showed dyadic behavioral trajectory similarity cluster membership to have a weak relationship with gender homophily.

The final model in each panel fits the two-way marginal between step and temporal degree cluster membership. As already noted, there is a good deal of association between these two cluster memberships. This is reflected in the substantial improvement in model fit. Furthermore, the residual deviance in these models is very low, indicating that a model with three-way interactions allowing the association between step and communication clusters to vary by homophily of the dyad would not yield much of an improvement in fit.

Tab. 1 – Fit-statistics of log-linear models of the association between step and communication trajectory similarity clusters

Resid. Df	*Resid. Dev*	*Df*	*Deviance*	*Pr(>Chi)*		*Eq.*
			Homophily: Gender			
471	4,984.68					S + C + G.H.
440	4,907.10	31	77.58	0.00	0.02	(S * G.H.) + (C * G.H.)
220	240.73	220	4,666.38	0.00	0.94	(S * G.H.) + (C * G.H.) + (S * C)
			Homophily: Race/Ethnicity			
465	6,320.90					S + C + R/E.H.
434	5,372.17	31	948.73	0.00	0.15	(S * R/E.H.) + (C * R/E.H.)
214	616.37	220	4,755.80	0.00	0.75	(S * R/E.H.) + (C * R/E.H.) + (S * C)
			Homophily: Religion			
468	6,286.57					S + C + Relig.H.
437	5,036.20	31	1,250.37	0.00	0.20	(S * Relig.H.) + (C * Relig.H.)
217	415.23	220	4,620.96	0.00	0.74	(S * Relig.H.) + (C * Relig.H.) + (S * C)

Note: Within each section, the p-value is for the comparison of one row with the preceding row. G.H. stands for gender homophily, R/E.H. stands for race/ethnicity homophily, and Relig. H. stands for religion homophily. S represents steps cluster, and C represents communication cluster.

6. Discussion

Dyads are an important component of social structure and, for some, the building block of social networks (Wasserman *et al.*, 1994, p. 505-ff; Rivera *et al.*, 2010). The analytic approach outlined here provides a way to take measures of time-varying traits defined on actors in a dynamic social network and quantify aspects of dyadic network evolution. This allows us to extend to the temporal case the basic notion of dyadic similarity; this notion is based on such basic social network constructs as homophily. Considering similarity to be a time-varying attribute of dyads yields the notion of a temporal *dyadic similarity trajectory*, which helps to specify the dynamic evolution of each pair of actors in the network concerning how similar (or dissimilar) they are on a target attribute. This attribute, as we have shown here, can be either exogenous or endogenous to the network. Building on this, we define the idea of *dyadic similarity trajectory cluster* as a mapping that assigns each dyad to a data-derived class based upon whether they share a temporal similarity pattern with other dyads in the same class.

Our empirical analysis, both in previous work and in this paper, shows that these clusters encode essential information, allowing us to predict both temporal network dynamics and whether a dyad is homophilous on a given set of (time-constant) traits. The results reported in this paper show that the dyadic trajectory similarity cluster approach previously shown to be fruitful when considering behavioral traits exogenous to the network (Sepulvado *et al.*, forthcoming) can be usefully extended to actor-level traits endogenous to the network, such as communicative activity (temporal degree). Our results indicate that dyadic trajectory similarity classes derived from this type of endogenous trait also encode useful information about network dynamics (e.g., helping us predict whether a dyad is connected or not) and are statistically associated with dyadic classes obtained from trajectory similarity based on exogenous traits. These findings indicate that insights into when and how social network ties form, how they evolve, and how long they persist can be garnered by constructing both behavioral and network position similarity trajectories for pairs of persons and ascertaining how these similarity patterns change when a tie is formed, during the life of the tie, before its decay, and after the tie no longer exists.

Future research should extend the approach proposed in this paper in several ways. First, temporal degree is only one of many endogenous time-varying traits defined on actors in a social network. Multiple indicators of an actor's position in a temporal network at a given time – defining a type of "temporal centrality" (Kim & Anderson, 2012) – yield an actor-level trajectory (Liu *et al.*, 2018), from which one may derive a corresponding dyadic similarity trajectory and a related cluster assignment for each dyad. Future work can thus investigate whether dyadic trajectory similarity clusters based on other endogenous measures of actors' position provide substantively relevant network information (e.g., helping predict other dyadic properties). Such work can also examine the link between these other trajectory clusters and other fine-grained actor-level traits.

Additionally, the approach proposed here can be extended to other network building blocks or "motifs" (Milo *et al.*, 2002) beyond dyads, such as triads and higher-level structures. After all, a dyad is a subgraph of size 2, and it is possible to extend the notions of similarity, similarity trajectory, and similarity trajectory clusters for subgraphs of larger size. For instance, triadic dissimilarity can be treated as an additive function of the dissimilarity between its three constituent dyads. The temporal evolution of this quantity thus gives triadic dissimilarity trajectories. Triads can then be assigned to triadic trajectory similarity clusters, and these could be used to help predict whether given triads belong to (or a more likely to transition into) well-known triadic connectivity classes (Wang *et al.*, 2014), such as the null, open, or closed triad.

References

Aghabozorgi S., Shirkhorshidi A.S., Wah T.Y. (2015), "Time-series clustering. A decade review", *Information System*, 53, pp. 16-38.

Bahulkar A., Boleslaw K.S., Chawla N., Lizardo O., Chan K. (2017), "Influence of personal preferences on link dynamics in social networks", *Complexity*, Article ID 4543563.

Bidart C., Lavenu D. (2005), "Evolutions of personal networks and life events", *Social Networks*, 27 (4), pp. 359-376.

Centola D. (2010), "The spread of behavior in an online social network experiment", *Science*, 329 (5996), pp. 1194-1197.

Christakis N.A., Fowler J.H. (2007), "The spread of obesity in a large social network over 32 years", *The New England Journal of Medicine*, 357 (4), pp. 370-379.

Christakis N.A., Fowler J.H. (2008), "The collective dynamics of smoking in a large social network", *New England Journal of Medicine*, 358, pp. 2249-2258.

Christakis N.A., Fowler J.H. (2013), "Social contagion theory: Examining dynamic social networks and human behavior", *Statistics in Medicine*, 32, pp. 556-577.

De Nooy W. (2011), "Networks of action and events over time. A multilevel discrete-time event history model for longitudinal network data", *Social Networks*, 33 (1), pp. 31-40.

Dickison M.E., Magnani M., Rossi L. (2016), *Multilayer social networks*, Cambridge University Press, Cambrifge.

Dokuka S., Valeeva D., Yudkevich M. (2016), "Homophily evolution in online networks: Who is a good friend and when?", *International Conference on Analysis of Images, Social Networks and Texts*, Springer, Cham, pp. 91-99.

Epskamp S. (2020), "Psychometric network models from time-series and panel data", *Psychometrika*, 85 (1), pp. 206-231.

Feld S.L. (1982), "Social Structural Determinants of Similarity among Associates", *American Sociological Review*, 47 (6),pp. 797-801.

Firth D. (1992), "Bias reduction of maximum likelihood estimates", *Biometrika*, 80 (1), pp. 27-36.

Fowler J.H., Christakis N.A. (2008), "Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study", *British Medical Journal*, 337, a2338.

Fowler J.H., Christakis N.A. (2008a), "The dynamic spread of happiness in a large social network", *British Medical Journal*, 337, a2338.

Fowler J.H., Christakis N.A. (2008b), "Estimating peer effects on health in social networks: A response to Cohen-Cole and fletcher; and Trogdon, Nonnemaker, and Pais", *Journal of Health Economics*, 27, pp. 1400-1405.

Holme P., Saramäki J. (2012), "Temporal networks", *Physics Reports*, 519 (3), pp. 97-125.

Ingram P., Morris M.W. (2007), "Do people mix at mixers? Structure, homophily, and the 'life of the party'", *Administrative Science Quarterly*, 52 (4), pp. 558-585.

Kim H., Anderson R. (2012), "Temporal node centrality in complex networks", *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 85 (2 Pt 2), 026107.

Kossinets G., Watts D.J. (2009), "Origins of homophily in an evolving social network", *American Journal of Sociology*, 115 (2), pp. 405-450.

Krackhardt D., Handcock M.S. (2006, June), "Heider vs Simmel: Emergent features in dynamic structures", in E. Airoldi, D.M. Blei, S.E. Fienberg, A. Goldenberg, E.P. Zing, A.X. Zheng (*eds.*), *ICML Workshop on Statistical Network Analysis*, Springer, Berlin, pp. 14-27.

Lazer D., Pentland A.S., Adamic L., Aral S., Barabasi A.L., Brewer D., Christakis N., Contractor N., Fowler J., Gutmann M. *et al.* (2009), "Life in the network: the coming age of computational social science", *Science*, 323 (5915), p. 721.

Leszczensky L., Pink S. (2019), "What drives ethnic homophily? A relational approach on how ethnic identification moderates preferences for same-ethnic friends", *American Sociological Review*, 84, 3, pp. 394-419.

Lewis K., Kaufman J. (2018), "The Conversion of Cultural Tastes into Social Network Ties", *The American Journal of Sociology*, 123 (6), pp. 1684-1742.

Lewis K., Kaufman J., Gonzalez M., Wimmer A., Christakis N. (2008), "Tastes, ties, and time: A new social network dataset using Facebook.com", *Social Networks*, 30 (4), pp. 330-342.

Liu S., Hachen D., Lizardo O., Poellabauer C., Striegel A., Milenković T. (2018), “Network analysis of the NetHealth data: exploring co-evolution of individuals’ social network positions and physical activities”, *Applied Network Science*, 3 (1), p. 45.

Martin J.L., Yeung K.-T. (2006), “Persistence of closer personal ties over a 12-year period”, *Social Networks*, 28, pp. 331-362.

McPherson M., Smith-Lovin L., Cook J. (2001), “Birds of a feather: Homophily in social networks”, *Annual Review of Sociology*, 27 (1), pp. 415-444.

Milo R., Shen-Orr S., Itzkovitz S., Kashtan N., Chklovskii D., Alon U. (2002), “Network motifs: simple building blocks of complex networks”, *Science*, 298 (5594), pp. 824-827.

Miritello G. (2013), *Temporal Patterns of Communication in Social Networks*, Springer Science & Business Media, Berlin.

Miritello G., Lara R., Cebrian M., Moro E. (2013), “Limited communication capacity unveils strategies for human interaction”, *Scientific Reports*, 3.

Morgan D.L., Neal M.B., Carder P. (1997), “The stability of core and peripheral networks over time”, *Social Networks*, 19 (1), pp. 9-25.

Paparrizos J., Gravano L. (2015), “k-Shape: Efficient and Accurate Clustering of Time Series”, in *SIGMOD ‘15: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, Victoria, Australia*, Association for Computing Machinery, New York, pp. 1855-1870.

Paparrizos J., Gravano L. (2017), “Fast and accurate time-series clustering”, *ACM Transactions on Intelligent Systems and Technology*, 42 (2), pp. 1-49.

Noel H., Nyhan B. (2011/7), “The ìunfriending’ problem: The consequences of homophily in friendship retention for causal estimates of social influence”, *Social Networks*, 33 (3), pp. 211-218.

Powers D., Xie Y. (2008), *Statistical Methods for Categorical Data Analysis*, Emerald Group Publishing, London.

Puhr R., Heinze G., Nold M., Lusa L., Geroldinger A. (2017), “Firth’s logistic regression with rare events: accurate effect estimates and predictions?”, *Statistics in Medicine*, 36, pp. 2302-2317.

Purta R., Mattingly S., Song L., Lizardo O., Hachen D., Poellabauer C., Striegel A. (2016), “Experiences measuring sleep and physical activity patterns across a large college cohort with fitbits”, *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, pp. 28-35.

Purta R., Striegel A. (2019), “Predicting Friendship Pairs from BLE Beacons Using Dining Hall Visits”, *2019 28th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1-9.

Raeder T., Lizardo O., Hachen D., Chawla N.V. (2011), “Predictors of short-term decay of cell phone contacts in a large scale communication network”, *Social Networks*, 33 (4), pp. 245-257.

Rivera M.T., Soderstrom S.B., Uzzi B. (2010), “Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms”, *Annual Review of Sociology*, 36 (1), pp. 91-115.

Schaefer D.R., Kreager D.A. (2020), “New on the Block: Analyzing Network Selection Trajectories in a Prison Treatment Program”, *American Sociological Review*, 85 (4), pp. 709-737.
Sekara V., Stopczynski A., Lehmann S. (2016), “Fundamental structures of dynamic social networks”, *Proceedings of the National Academy of Sciences of the United States of America*, 113 (36), pp. 9977-9982.
Sepulvado B., Wood M.L., Fridmanski E., Wang C., Chandler C.J., Lizardo O., Hachen D. (2020), “Predicting Homophily and Social Network Connectivity From Dyadic Behavioral Similarity Trajectory Clusters”, *Social Science Computer Review*, forthcoming.
Wang C., Lizardo O., Hachen D.S. (2014), “Triadic evolution in a large-scale mobile phone network”, *Journal of Complex Networks*, 3 (2), pp. 264–290.
Wang C., Lizardo O., Hachen D.S. (2020), “Neither influence nor selection: Examining co-evolution of political orientation and social networks in the NetSense and NetHealth studies”, *PloS One*, 15, 5, e0233458.
Wasserman S., Faust K. (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.
Yang Y., Nitesh V.C., Uzzi B. (2019), “A network’s gender composition and communication pattern predict women’s leadership success”, *Proceedings of the National Academy of Sciences*, 116, 6, pp. 2033-2038.

3. *Do you influence me? Evidence from a case study of network ties among university students in Pisa*

by Anna Piazza[*], Srinidhi Vasudevan[**]

1. Introduction

Peers have a great influence on how students behave, on their own educational outcomes and what career paths they will take (for recent reviews see Sacerdote, 2014). Social influence occurs when the behaviour of individuals is affected by social interactions with other peers. As such, individuals change their behaviour in response to their relationships with their peers (Rambaran *et al.*, 2017) whereby individuals weigh and combine their own and others' behaviour before revising their own (Friedkin & Johnsen, 2011). Extant research on peer influence in education setting has focused on whether student outcomes are influenced by the composition of their peers, such as gender, race and ability (Griffith & Main, 2019; Cools *et al.*, 2019); family support (Cheng, Ickes & Verhofstadt, 2012), and co-location such as being roommates or dormmates (Frijters *et al.*, 2019, Sacerdote, 2011; Parker *et al.*, 2010). For example, it is seen that the grades of individual students are influenced by their peers who are enrolled in the same course at the University (Brunello, De Paola & Scoppa, 2010). However, results based only on the properties of peers do not tell the complete story about the existence of peers influences because the literature suggests "that we do not yet know enough about the nature of peer effects" (Sacerdote, 2014, p. 269).

The existing studies have looked at how exposure to a general influence impact student outcome and do not consider the interpersonal influ-

[*] Department of Science, Technology, Engineering and Public Policy, University College London, a.piazza@ucl.ac.uk; Department of System Management and Strategy, University of Greenwich, anna.piazza@greenwich.ac.uk.

[**] Department of Science, Technology, Engineering and Public Policy, University College London, s.vasudevan@ucl.ac.uk.

ence which examines how an individual's social network influences their outcomes. Given the social nature of peers in educational settings, available research suggests that network of interactions between students are important channels for the diffusion of the ongoing influence process (Berthelon *et al.*, 2019; Bramoullé *et al.*, 2009). In the context of the continuing debate on whether student outcomes are influenced by the network of their peers (Bond *et al.*, 2017; Patacchini and Zenou, 2016; Calvó-Armengol *et al.*, 2009), we study the role of social networks on student outcome, i.e. performance in terms of grades and satisfaction at university level. What this paper adds to the established literature is a focus on the mechanism that may be responsible for transmitting social influence. By doing this, we contribute to the few studies that examine the role of social networks at the university level (Stadtfeld *et al.*, 2019; Torlò & Lomi, 2017; Vitale, Porzio & Doreian, 2016; Lomi *et al.*, 2011).

The aim of this paper is therefore to empirically assess individual's social network influence on student's outcomes using a sample of a cohort post-graduate students who attend the same course at Pisa University in the academic year of 2019-2020. Our setting is exceptional because, even though educational settings have an important role for the development of socialisation among students and individual outcomes (Lomi *et al.*, 2011; Jensen & Parsons, 1959), this particular setting does not offer social space where students can socialise. The following statement illustrates the exception of our setting: "We must start from the concrete observation that students who attend the courses at the Polo Piagge (that is the building where Political Science and Law lectures take place) do not have the opportunity to participate in events, occasions and common spaces (that are examples of environments socialisation among students is frequently encouraged). In general, the way the built environment and teaching/learning are organized around individual study oriented; that can be certainly effective, however, it does not promote socialisation where teaching and learning are organised around a well-equipped study room to study together in a group; occasions for meeting, exchange knowledge and information on course-related matters"[1].

It is hence well-suited to address questions that are important in explaining and quantifying the social nature of peers: Does socialisation affect student's performance? Does influence process emerge only through the social compositions of these peers or is there evidence that influence emerge through social relationships among peers? Do advice networks influence stu-

[1] This statement is from the Professor of the Department of Political Science at the Pisa University.

dents' performance and satisfaction? Does built environment affect how students socialise? Our research offers preliminary answers to these questions and might be used as a first evaluation for quantifying the peers' influence in this sample. These questions are not only theoretically relevant to academic scholars but also significant for policy makers and educators. If students' performance is affected by ongoing social relations, policies aimed at understanding the process of socialisation among students may be important strategies for increasing student potential achievement at University level.

For many scholars, the estimation of the social influence hypothesis and therefore the effect of peers has been done by using a linear-in-means models (Sacerdote, 2011; Bramoullé, Djebbari & Fortin, 2009). In this model, the individual's outcome is shaped not only by their own characteristics but also by the average characteristics of the group of individuals they are associated with. They assume that individuals in a network are embedded in sub-groups and are not linked to members outside of this sub-group (Vitale, Porzio & Doreian, 2016). As Vitale *et al.* (2016) argue: "Linear-in-mean models are not really well suited for the study of the presence of social influence (or peer effects) on student performance at the university level, and suggest other statistical models be adopted". This alternative statistical model is the network autocorrelation model (Doreian, 1989; Doreian, Teuter & Wang, 1984; Dow, Burton & White, 1982) that enables to capture and/or control for social endogeneity in the existence of non-exogenous covariates which is a consequence of the interactions between individuals. In other words, the network autocorrelation model allows to quantify the social process that are contained within the networks pose methodological complications due to the dependencies in the data, while the sociological covariates such as individual characteristics and interaction group effects are entered into the model (Vitale, Porzio & Doreian, 2016). Therefore, we adopt the network autocorrelation model while controlling for sociological and other covariates. We find that the advice network among students positively influences their performance and their satisfaction and, in both cases, other individual covariates such as student background, gender does not have an impact. Additionally, this paper also controls for the building effects in understanding student performance and satisfaction. We find that the built environment is positive and significant in explaining the performance and satisfaction.

The paper is organised as follows. The theory is developed in section 2. Empirical setting, data and methodology are described in section 3. The estimation of the empirical model and the results is described in section 4. The conclusion and potential directions for future research are discussed in section 5.

2. Influence and relations on student performance

Research on social influence reports that the key benefit of social relations is the possibility to access information that are channelled through networks e.g. social capital (Coleman, 1988) which increases the chance of changing an individual's performance (Carbonaro & Workman, 2016). This happens because the social networks provide the relational structure by which many processes occur, such as social learning (Mason, Conrey & Smith, 2007); social integration and collaboration (Stadtfeld *et al.*, 2019), assimilation (Lomi *et al.*, 2011). The role of social networks and its relationship with student outcomes is acknowledged in the study in primary and secondary schools. For example, Bond *et al.* (2017) examine the relationship between individual networks and achievement and find that students that are embedded in low-achievement network are more likely to decrease their own achievement. In addition, they show that students who occupy a central position within the network are more likely to increase their achievement. In a similar way, using a sub-sample of the same data Calvó-Armengol *et al.* (2009) estimate social influence process through the structure of individual networks to explain student outcomes. They find that the network position has an impact on individual outcomes. They show that the network position accounts for 7% of the increase in the standard deviation of individual academic performances. Lavy *et al.* (2012) determine the impact of networks on student outcomes and they find that student achievement varies according with low-achieving social networks and high-achieving social networks. This body of work examines the effect of network of interactions on student performance because of the significant relevance of the peers in the lives of the students. However, the appropriate bases of social influence also depend on the specific context. As such, our work contributes to the existing literature of social influence study and directly examines the social influence process in the university context.

2.1. Formalising research hypotheses

In the university, students are more likely to interact more where having relationships with each other influence positively their own performance (Vitale, Porzio & Doreian, 2016). Social influence with regards to performance can occur as a result of one student having network ties with another student who is knowledgeable and/or a high achiever. Students are more likely to achieve higher levels of academic performance as a consequence of being influenced by their peers who provide knowledge on specific topics or advice on how to perform better (Frank *et al.*, 2008). Indeed, individual academic performance is shaped by the existence of underlying ongoing advice relations supporting the transfer of knowledge (Lomi *et al.*, 2011). It is also possible that a student can be motivated to work harder, and hence be satisfied, because of having relationship with a high performing student. Consistent with this mechanism, we formulate the following hypotheses:

- **H1 (a)** the higher is the interactions among students, namely stronger levels of social influence, the higher is the tendency to perform better;
- **H2 (b)** the higher is the interactions among students, namely stronger levels of social influence, the higher is the tendency to be satisfied.

Scholars (Pasalar, 2007; Schneider, 2002) acknowledge that student performance is also influenced by the built environment that promotes socialisation processes within the school (Osher *et al.*, 2014). Research shows a positive correlation between the building environment and the socialisation process whereby students are more likely to interact more with each other due to the design of the space (Moore & Lackney, 1993). A study by Garibaldi and Josias (2015) offers a persuasive evidence on the role of the built environment on academic performance whereby space facilitates socialisation process. The importance of the space is given by its social construct shaped by social aspects of the environment. One social aspect of the environment is connected to the social conditions for learning, where students tend to be cooperative in teamwork and have a sense of belonging. "Instruction and learning are optimized with flexible spaces, such as classroom with designed centers where students decide what, where and how they will learn, enabling student interactions" (Garibaldi & Josias, 2015, p. 1591). Research shows that the space of the building like common spaces (or public spaces), where easy access is available to students, is positively associated with high levels of social interactions among students (Pasalar, 2003). Space might reflect the establishment of regular routines encouraging socilisation among students. However, this line of research is still limited. Determining the role of the built environment is empirically important to understand how the space impacts the academic performance and the ongoing social influence (socialisation) in a specific context. Indeed, our research directly captures the importance of the building environment by looking at the space. Given this, our second hypothesis is:

- **H2 (a)** greater the socialisation spaces that the built environment offers, the higher is academic performance;
- **H3 (b)** greater the socialisation spaces that the built environment offers, the higher is satisfaction.

Consistent with prior research (Sacerdote, 2011), we assume and control that other sociological covariates, like shared individual characteristics that might affect student academic performance.

3. The case study

3.1. Setting, data and methodology

Educational settings are both testing grounds for individual achievement as well as agents of socialisation (Akerlof & Kranton, 2002), therefore they represent an ideal setting for developing and accessing research hypotheses about social influence. We collected data on 63 students involved in the Social sciences field enrolled in a master's level program at a Pisa University in the academic year of 2019-2020. The master level program attracts students oriented toward careers in social work or management of social institutions, public and private enterprises. Students enrolled in the course were not assigned into specific groups therefore they were willing to share knowledge on specific topics and also about how they have performed academically. Indeed, achievement in terms of grades are seen as one of the determinants for advice relationships to exist (Snijders & Lomi, 2019).

The class consisted of a group of students who came from different Italian regions with similar academic backgrounds and a variety of prior work experiences. The data were collected during the second academic term in the second-year graduate program. The students completed the survey instrument as part of their academic session. The survey instrument included sections on demographics and network questions. The survey was administered through email; where the questionnaire was reviewed and approved by the Statistical Office of Pisa University. Informed consent was obtained from participants[2]. Only two students did not participate in the survey; our final dataset was 61 and therefore the response rate was empirically good 96.8%. In our final sample, the age range is 22-29 years (mean = 26.73) with 85% percent women. 61% of the students came from different Italian regions within the country but lived in Pisa, 23% commuted from other regions to Pisa and the 16% were locals who lived in Pisa. 82% of the students have attended a lyceum at secondary school.

[2] In line with Research Ethics regulation.

3.2. Variables and measures

We use two dependent variables to capture academic outcomes, namely performance and satisfaction.

Performance: performance was measured by using an average score of the students for all the exams taken by the student in the year. A score of 0 represented that the examinations were not taken by the students.

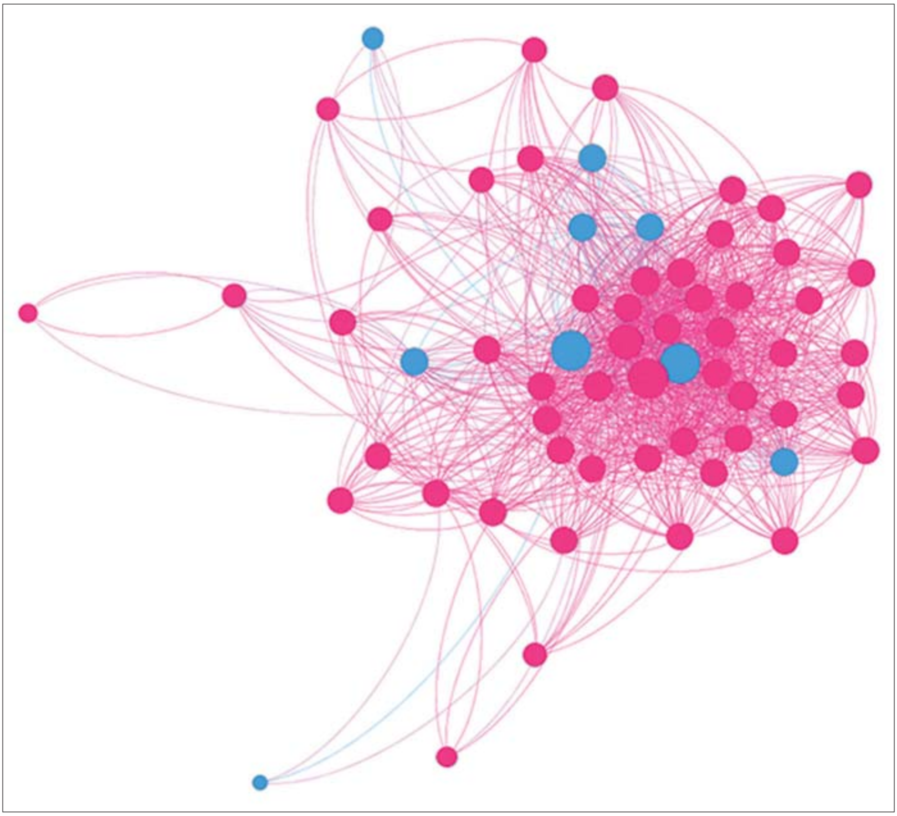
Satisfaction: satisfaction was measured by using an average score of nine items that were captured in the questionnaire (such as quality of learning materials, quality of relationship with tutors, quality of relationship with classmates, the personal and professional development). The reliability of these items were measured using the Cronbach Alpha which was 0.85. Satisfaction was captured to understand how students perceived they achieved their cognitive learning outcomes. Satisfaction is seen as a proxy of performance and is captured by several universities globally. This is measured to see if students believe that they have achieved the learning outcomes which ultimately leads to better performance (for review, see Bedggood & Donovan, 2012).

3.2.1. Network measures

Advice network: we collected social networks data through the well-established roster method (Marsden, 2011). Each student was asked to answer the following question: "Please indicate the names of your classmates to whom you would go for advice on course-related matters. Please rate on five-point scale the extent to which you seek advice with 1 being very rarely and 5 being almost always". The respondents were then presented with a list of names in alphabetical order (by last name) and could check the box next to the individual if they considered them a person they went to for advice and indicate the five-point scale. Initially, we have a valued-matrix, however, for the purposes of this study, we dichotomised the network with a threshold value of greater than 1 because we wanted to capture all advice-seeking behaviour (to test the social influence hypothesis) within the class and the intensity of the relationships was not the aim of the current study.

Figure 1 shows a graphical representation of the resulting advice network.

Fig. 1 – Advice Network



Each point (or node) in the graph represents a student. The colour of the node represents gender (pink: female, blue: male). Lines between nodes represent an advice relationship. The network is binary and directed. Node size is proportional to academic performance with bigger node size representing better performance. Table 1 reports some network statistics of the advice network. The network density shows high levels of interactions among students. Students on an average ask 20 of their classmates for advice. The standard deviation for out-degree is 13.333 and the standard deviation for in-degree is 12.498. This suggests that while students commonly seek advice, it is less common to be sought for advice by classmates. The average path length and network diameter are small and shows that students can connect with others with minimal steps.

Tab. 1 – Network Descriptive Statistics

Network density	0.345
Average degree	20.705
Standard deviation of outdegree	13.333
Standard deviation of indegree	12.498
Average Path length	1.74
Network Diameter	4

3.2.2. Control variables

There are several alternative explanations for individual student performance that we control for. These include the building environment effect. This 5-item Likert scale was constructed for the present study in order to capture the space and layout offered by the building understudy. Each items consisted of statements on the space offered (or not) by the building. The respondents have to choose one item. We ask to the respondents why the space and the layout are (not) designed to promote socialisation among students. We also control for age that is in number of years; for gender, for parent's education (parents' background as number of years in higher education). We control for whether a student mainly spend all his/her time to study and or to work. The students are separated in three categories, coded I do not work, but I only study = 1; I mainly work and study when I have time = 2; I mainly study and work when I have time = 3. This variable is context-depending due the particular segmentation of the Italian job-market. In Table 2 are the variable descriptive statistics that are used for this study.

Tab. 2 – Variables

Variable	*Definition*	*Type*	*Average (St.Dev.)*
Gender	Gender of the student	Categorical (dummy)	85% female
Enrolment age	Age of the student	Continuous	26.73 (6.62)
Living near the university location	To capture students travelling from other destinations to Pisa, local students and students from other destinations living in Pisa followed the Diritto allo Studio Universitario (DSU) classification	Categorical (dichotomous dummy with a reference category)	61% from different Italian regions but lived in Pisa; 23% commuted; 16% were locals who lived in Pisa
Parents' background	Education of parents captured in terms of years after high school	Continuous	3.64 (0.79)
Student background	The capture the background of the student to refer to "Istituto tecnico", "Classic Lyceum" and "Social Science Lyceum"	Categorical (dichotomous dummy with a reference category)	40.98% Istituto tecnico, 40.98% liceo classico, 18.03% liceo scienze sociali
Bachelors grade	The grade received by students in their bachelor's degree to include three categories	Categorical (dichotomous dummy with a reference category)	49% with a grade of Da 100 a 107, 18% with a grade of Da 108 a 110L, 31% with a grade of Fino a 99
Building effect	To capture the space offered for the students	Categorical (Likert-scale 1-5, polytomous)	26.22% with value 1, 52.45% with value 2, 4.91% with value 3, 6.55% with value 4 and 9.83% with value 5
Student satisfaction	To capture the perception of students feeling satisfied with achieving their cognitive learning outcomes	Continuous	7.24 (1.27)
Average grade	The average of all exams taken by the student in the given year	Continuous	23.36 (9.34)
Student status	To capture whether the student was a full-time student (Category1), worked part-time and spent some time studying (Category 2) or studied most of the time and worked only for some hours part-time (Category 3)	Categorical (dichotomous dummy with a reference category)	21.31% in Category 1, 16.39% in Category 2, 34.43% in Category 3

3.3. Methodology

Linear Network Autocorrelation models allow for quantifying the influence of peers in a network while taking into account node-related characteristics (in our case, gender, students' background, family background, prior grades). The network correlation effect or the network effect ρ identifies the presence of and the magnitude of the peer influence. The researcher uses the inferential test to seek evidence for social influence present using the null hypothesis. When H0: ρ is rejected, there is evidence that there is some degree of influence present in the network.

Autocorrelation models allow the researcher to test social influence effects through the specification which allows to test specific theories and alter the specification dependent on the theory. We use Network autocorrelation model (NAM) to evaluate social influence due to the presence of interdependent units which are embedded in the network. This interdependency will violate the regression assumptions for obtaining coefficient estimates that are unbiased. Within the NAM, there are two classes of models that help estimate the social influence, namely the network disturbances model and the network effects model. In the latter model, the autocorrelation parameter is included in the dependent variable y to account for the interdependencies in the model. Network effects model is used in this research as it provides an opportunity to estimate the strength of the social influence through the parameter ρ (Vitale *et al.*, 2016; Doreian, 1989).

4. Model specification and results

The general form of the network effects model is:

$$Y = \rho Wy + X\beta + \varepsilon \tag{1}$$

where, the assumption is that the error term is distributed normally.

In the specification (1), y represents the endogenous dependent variable consisting of n values for the individuals in the network. X represents the covariate matrix for p covariates and n individuals (nXp). This also includes the intercept term's unit vector. The network matrix is represented by W which is an n X n matrix with elements wij which is used to measure how actor i is influenced by actor j and is given by ρ.

This model has been chosen for two reasons. Firstly, this enables to model the academic performance of the individual as a function of their prior education, socio-demographic attributes as well as the performance of the network neighbours. Secondly, the ρ parameter can be used to understand the network effect's strength.

We use two dependent variables namely performance and satisfaction. The full model is reported in Equation 2 and 3.

Performance =

$$\alpha + pWAdviceNetwork + \beta_1 Gender + \beta_2 Age1 + \beta_3 Age2 + \beta_4 Location1 + \beta_5 Location2 + \beta_6 StudentBackground + \beta_7 ParentsBackground + \beta_8 BachelorsGrade1 + \beta_9 BachelorsGrade2 + \beta_{10} StudentStatus1 + \beta_{11} StudentStatus2 + \beta_{12} BuildingEffect + \varepsilon \tag{2}$$

Satisfaction =

$$\alpha + pWAdviceNetwork + \beta_1 Gender + \beta_2 Age1 + \beta_3 Age2 + \beta_4 Location1 + \beta_5 Location2 + \beta_6 StudentBackground + \beta_7 ParentsBackground + \beta_8 BachelorsGrade1 + \beta_9 BachelorsGrade2 + \beta_{10} StudentStatus1 + \beta_{11} StudentStatus2 + \beta_{12} BuildingEffect + \varepsilon \tag{3}$$

5. Results

We estimate the influence effect of the advice network and Model 1 (Equation 2) uses the dependent variable performance and Model 2 (Equation 3) uses the dependent variable satisfaction. We use R 3.6.1 and the libraries SNA, iGraph (for plotting the networks) and lnam package for estimating the network autocorrelation model.

In both models, the estimates provide evidence that social influence acts through advice relations. More specifically, in model 1 the network is positive and significant meaning that students are more likely to perform better as they are influenced by their advisors. The estimated parameter for built environment is strongly significant and shows that when the built environment facilitates the interaction among students, they tend to perform better. The estimated parameters for the control variables are generally statistically weak. Studentstatus2, is negative and significant to student performance and this indicates that students who work part-time and study part-time have a lower performance compared to the other two categories of students who study full-time and those that study most of the time and work part-time occasionally. Students whose parents have higher qualifications (parents' background) have a higher likelihood of performing better.

Tab. 3 – Results of network autocorrelation model – Standard error in parentheses

Coefficients	*Model 1 (performance)*	*Model 2 (satisfaction)*
	-2.5507	-0.3239
	(2.8106)	(0.5209)
Age1	2.1996	1.1876
	(3.2245)	(0.6020)**
Age2	3.0121	2.1577
	(2.8181)	(0.5265)***
Location1	3.5479	0.0214
	(3.0220)	(0.5646)
Location2	-0.2980	-0.4987
	(2.4570)	(0.6483)
Student Background	0.9280	0.6955
	(2.4570)	(0.4587)
Parents Background	3.1950	0.8221
	(1.1724)**	(0.2197)***
BachelorsGrade1	-3.3710	0.6203
	(3.1677)	(0.2953)
BachelorsGrade2	0.3571	0.7076
	(2.7931)	(0.1754)
StudentStatus1	-2.9494	0.2334
	(2.8861)	(0.5339)
StudentStatus2	-10.7043	0.4862
	(3.6075)**	(0.6744)
Building Effect	4.2728	0.3467
	(1.0740)***	(0.2011).
Social Influence	0.0076	0.0067
	(0.0033)*	(0.0020)***
AIC	451.9	247.1
BIC	481.5	276.6

Signif. codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 "" " 1

Note: The AIC and BIC are for the estimated models. For each relation, the magnitude of the network effect is captured by the parameter ρ.

In Model 2, the advice network is positive and strongly significant in explaining satisfaction of students showing that students who interact more have a perception of achieving the desired learning outcomes better. The estimated parameter for the built environment is strongly significant and shows that socialisation process is fostered through the spaces offered and this is evidenced through a positive impact on satisfaction of individuals. The estimated parameters for the control variables are generally statistically not strong. Age affects individual satisfaction and those who are between 22-25 years old (captured through variables age1 and age2) have a higher satisfaction compared to those who are over 26 years. Students whose parents have higher qualifications (captured through parents' background) have higher satisfaction.

6. Discussions and conclusion

What happens in university lays the groundwork for work careers, but students may not always have such a long-term view. What may be important to them is what is immediate – in a short period –, such as successfully passing a course through knowledge and information sharing. In this study, we have shown that when students attend the same course their individual task performance is influenced by their peers through advice relations. As a consequence, peers' influence contributes to the fundamental process of interactions among students that affect educational achievement and satisfaction. We extended on previous research by using NAM to assess influence process at university setting. Moreover, we investigated the predictions that build environment – specifically the space that is offered – impacts on students socialisation and therefore their outcomes.

This study yielded detailed insights into the mechanism of social influence as an important point in the formative process of interactions among academic student at university level. Understanding the interaction is important as learning and development is a social activity (Egan, 1980) and this facilitates learners to have a meaningful dialogue thereby enabling them to become satisfied learners. At a master's level this becomes even more important as the goal of the courses is not to see if students can repeat or memorise but to ensure they are able to exhibit, demonstrate contextual knowledge (Brooks & Brooks, 1993, p. 16). Prior qualitative studies at school and undergraduate levels show that students believed that when they had a chance to have social interactions, they have been able to be more interested in learning and achieve their goals faster and in a more efficient manner (Vacca, Mraz & Vacca, 2017; Hurst, Wallace & Nixon, 2013). This study also contributes to understanding how the built environment impacts socialisation especially at a university setting. While the built environment intends to offer social spaces at universities, the design often does not reflect this message. The environment is devoid of communal activities as the building does not provide social spaces. When the social spaces at the university are well thought-out, this results in fruitful encounters and interactions among students. Over time, studies show that getting the spatial aspects and the built environment wrong at the university would result in students losing focus and concentration (Turner, Scott-Young & Holdsworth, 2016). This study contributes to furthering the understanding between the built environment and subsequently student performance by unequivocally showing that better interaction that is facilitated by the built environment acts as a backdrop to learning and leads to better performance and satisfaction.

By bringing these two main aspects to light, our work contributes to furthering reflection on the importance of social relationships and its mediating influence on student's outcomes. We are aware, however, that network effects are likely to be context-dependent, but the specific aspect of student outcome may be of interest in any one situation. We focus on one specific relation, advice relation between students. However, advice relation and knowledge sharing are not the only form of interactions. We choose this relation because prior studies have shown its importance and how the exchange of knowledge encourages social learning by promoting the exchanging of information (Mason, Conrey & Smith, 2007) in the education setting. In our work this is important because social learning is the main micro-process underlying the influence of student's performance. We have focused on social learning as one of possible effect induced by the exchange of information, and in general socialisation, that is social learning, because we are interested in understanding the social mechanism correlated with student's performance. Clearly, the establishment of social relationships and its micro-processes are different such as social integration, cohesion, collaboration, but the results of this work extend the understanding of social learning to the extent the motivation that encourages socialisation among students is exchanging knowledge and information on course-related matters.

While this study contributes to the existing literature on social influence and student performance the main limitation is the research design, cross-sectional. It would be useful to perform additional tests, more specifically by using longitudinal data, which would take into account the dynamic nature of the time and the network (Stadfeld *et al.*, 2019). Collecting longitudinal data help understand the dynamics in the group structures within the network to distinguish the mechanisms of selection and influence (Shalizi & Thomas, 2011). A longitudinal study would strengthen our claims, as it would enable us to control for sources of unobserved heterogeneity among individuals, however, given the difficulties in reaching out to this population due to students finishing their course, we have used a cross-sectional design. Secondly, we only capture the effect of the social space that is while considering the building effect. We are aware that other several factors such as aesthetics, layout can capture this, however we have only considered space as a relevant factor for this research as it fosters the socialisation process and creates conducive environments for interaction and learning (McGregor, 2004).

Nonetheless, our study has a pivotal practical implication within the university context. The findings have policy implications as universities can take into account the socialisation process among students and how this is facilitated and include these in the strategies that aim to improve student performance and satisfaction.

References

Akerlof G.A., Kranton R.E. (2002), "Identity and schooling: Some lessons for the economics of education", *Journal of Economic Literature*, 40 (4), pp. 1167-1201.

Bedggood R.E., Donovan J.D. (2012), "University performance evaluations: what are we really measuring?", *Studies in Higher Education*, 37 (7), pp. 825-842.

Berthelon M., Bettinger E., Kruger D.I., Montecinos-Pearce A. (2019), "The structure of peers: the impact of peer networks on academic achievement", *Research in Higher Education*, 60 (7), pp. 931-959.

Bond R.M., Chykina V., Jones J.J. (2017), "Social network effects on academic achievement", *The Social Science Journal*, 54 (4), pp. 438-449.

Bramoullé Y., Djebbari H., Fortin B. (2009), "Identification of peer effects through social networks", *Journal of Econometrics*, 150 (1), pp. 41-55.

Brooks J.G., Brooks M.G. (1993), *Structuring learning around primary concepts: The quest for essence. In search of understanding: The case for constructivist classrooms*, Association for Supervision and Curriculum Development, Alexandria (VA).

Brunello G., De Paola M., Scoppa V. (2010), "Peer effects in higher education: Does the field of study matter?", *Economic Inquiry*, 48 (3), pp. 621-634.

Calvó-Armengol A., Patacchini E., Zenou Y. (2009), "Peer effects and social networks in education", *The Review of Economic Studies*, 76 (4), pp. 1239-1267.

Carbonaro W., Workman J. (2016), "Intermediate peer contexts and educational outcomes: Do the friends of students' friends matter?", *Social Science Research*, 58, pp. 184-197.

Cheng W., Ickes W., Verhofstadt L. (2012), "How is family support related to students' GPA scores? A longitudinal study", *Higher Education*, 64 (3), pp. 399-420.

Coleman J. (1988), "Social capital in the creation of human capital", *American Journal of Sociology*, 94, S95-S120.

Cools A., Fernández R., Patacchini E. (2019), *Girls, boys, and high achievers* (No. w25763), National Bureau of Economic Research, National Bureau of Economic Research, Cambridge (MA).

Doreian P. (1989), “Network autocorrelation models: Problems and prospects”, *Spatial statistics: Past, present, future*, Michigan Document Services Ann Arbor, pp. 369-389.
Doreian P., Teuter K., Wang C.-H. (1984), “Network autocorrelation models: Some Monte Carlo results”, *Sociological Methods & Research*, 13 (2), pp. 155-200.
Dow M.M., Burton M.L., White D.R. (1982), “Network autocorrelation: a simulation study of a foundational problem in regression and survey research”, *Social Networks*, 4 (2), pp. 169-200.
Egan K. (1980), “John Dewey and the social studies curriculum”, *Theory & Research in Social Education*, 8 (2), pp. 37-55.
Frank K.A. *et al.* (2008), “The social dynamics of mathematics coursetaking in high school”, *American Journal of Sociology*, 113 (6), pp. 1645-1696.
Friedkin N.E., Johnsen E.C. (2011), *Social influence network theory: A sociological examination of small group dynamics*, Cambridge University Press, New York.
Frijters P., Islam A., Pakrashi D. (2019), “Heterogeneity in peer effects in random dormitory assignment in a developing country”, *Journal of Economic Behavior & Organization*, 163, pp. 117-134.
Garibaldi M., Josias L. (2015), “Designing schools to support socialization processes of students”, *Procedia Manufacturing*, 3, pp. 1587-1594.
Griffith A.L., Main J.B. (2019), “First impressions in the classroom: How do class characteristics affect student grades and majors?”, *Economics of Education Review*, 69, pp. 125-137.
Hurst B., Wallace R., Nixon S.B. (2013), “The impact of social interaction on student learning”, *Reading Horizons: A Journal of Literacy and Language Arts*, 52 (4), p. 5.
Jensen G., Parsons T. (1959), “Chapter III: The Structure and Dynamics of Classroom Groups and Educational Systems”, *Review of Educational Research*, 29 (4), pp. 344-356.
Lavy V., Silva O., Weinhardt F. (2012), “The good, the bad, and the average: Evidence on ability peer effects in schools”, *Journal of Labor Economics*, 30, pp. 367-414.
Lomi A., Pallotti F. (2012), “Relational collaboration among spatial multipoint competitors”, *Social Networks*, 34 (1), pp. 101-111.
Lomi A., Snijders T.A., Steglich C.E., Torló V.J. (2011), “Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance”, *Social Science Research*, 40 (6), pp. 1506-1520.
Marsden P.V (2011), “Survey methods for network data”, *The SAGE Handbook of Social Network Analysis*, 25, pp. 370-388.
Mason W.A., Conrey F.R., Smith E.R. (2007), “Situating social influence processes: Dynamic, multidirectional flows of influence within social networks”, *Personality and Social Psychology Review*, 11 (3), pp. 279-300.
McGregor J. (2004), “Spatiality and the place of the material in schools”, *Pedagogy, Culture and Society*, 12 (3), pp. 347-372.

Moore G.T., Lackney J.A. (1993), "School design: Crisis, educational performance and design applications", *Children's Environments*, JSTOR, pp. 99-112.
Osher D., Kendziora K., Spier E., Garibaldi M.L. (2014), "School influences on child and youth development", in *Defining prevention science*, Springer, Boston, pp. 151-169.
Parker J. *et al.* (2010), *Classmate peer effects: Evidence from core courses at three colleges*, Reed College, Portland.
Pasalar C. (2003), *The effects of spatial layouts on students' interactions in middle schools: Multiple case analysis*, unpublished doctoral dissertation, North Carolina State University.
Pasalar C. (2007), "Spaces for Learning Through Better Social Interaction", in *School Building Design and Learning Performance: with a Focus on Schools in Developing Countries*, Comportements and Authors, Lausanne, pp. 51-60.
Patacchini E., Zenou Y. (2016), "Social networks and parental behavior in the intergenerational transmission of religion", *Quantitative Economics*, 7 (3), pp. 969-995.
Rambaran J.A. *et al.* (2017), "Academic functioning and peer influences: A short-term longitudinal study of network-behavior dynamics in middle adolescence", *Child Development*, 88 (2), pp. 523-543.
Sacerdote B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?", in E. Hanushek, S. Machin, L. Woessmann (*eds.*), *Handbook of the Economics of Education*, Elsevier, North-Holland, ch. 4, pp. 249-277.
Sacerdote B. (2014), "Experimental and quasi-experimental analysis of peer effects: two steps forward?", *Annu. Rev. Econ. Annual Reviews*, 6 (1), pp. 253-272.
Schneider M. (2002), "Do School Facilities Affect Academic Outcomes?", ERIC Publications, Washington (DC).
Shalizi C.R., Thomas A.C. (2011), "Homophily and contagion are generically confounded in observational social network studies", *Sociological Methods & Research*, 40 (2), pp. 211-239.
Stadtfeld C. *et al.* (2019), "Integration in emerging social networks explains academic failure and success", *Proceedings of the National Academy of Sciences*, 116 (3), pp. 792-797.
Torlò V.J., Lomi A. (2017), "The network dynamics of status: assimilation and selection", *Social Forces*, 96 (1), pp. 389-422.
Turner M., Scott-Young C., Holdsworth S. (2016), *Bouncing back to move forward: Resilience of students in the built environment*, Association of Researchers in Construction Management, Melbourne, 1, pp. 589-598.
Vacca R.T., Mraz M., Vacca J.A.L. (2017), *Content area reading: literacy and learning across the curriculum*, Pearson Education, New York.
Vitale M.P., Porzio G.C., Doreian P. (2016), "Examining the effect of social influence on student performance through network autocorrelation models", *Journal of Applied Statistics*, 43 (1), pp. 115-127.

4. *A structural analysis proposal for inter-organizational networks: doing a multilevel analysis through Multilevel ERGMs*

by Antonietta Riccardo*, Irene Psaroudakis**

1. Introduction

The objective of this paper is to investigate the interdependence between two inter-organizational networks of an Italian welfare community and their community-based tasks from a multilevel analysis perspective, using MERGs techniques.

This discussion is part of a more complex study on two Third Sector networks locally operating in Southern Italy, started in 2017 and conducted by a research group[1] affiliated to University of Sannio and University of Pisa. We develop here the results gained about the generative networks promoted by the Caritas centres operating in local territories corresponding to two dioceses of Aversa (RCA) and Benevento (RCB), as presented in the article "Inter-Organizational Networks and Third Sector: Emerging Features from Two Case Studies in Southern Italy" (Salvini *et al.*, 2020), and we underline the importance of setting up networking practices in the light of the Third Sector Code precepts, by enhancing social capital of territories and community-building dynamics. The Caritas dioceses networks realize the purpose of generating inter-organizational networks, promoting virtuous exchange projects, social initiatives, and collaborations, involving firstly Third Sector entities (voluntary organizations, cooperatives, associations, second-level organizations), but also local governments (most of all social-health care institutions), educational institutions (schools, universities), and church or-

* Department of Political Sciences, University of Pisa, antonietta.riccardo@phd.unipi.it.

** Department of Political Sciences, University of Pisa, irene.psaroudakis@sp.unipi.it.

1 The research group is composed by Andrea Salvini, Francesco Vasca, Antonietta Riccardo, Irene Psaroudakis.

ganizations, in a generative dynamic animating local communities. By this, they are considered as inter-organizational networks as they "intentionally created relationships among at least three autonomous but interdependent organizations that cooperate in order to achieve a common result (output) or to jointly produce an expected emergent behaviour (*outcome*)" (ivi, p. 210).

The study about the Third Sector inter-organizational networks in Southern Italy has been structured in two phases, and its aim is twofold: researchers wanted to demonstrate a) how metric properties of the networks can be useful indicators to monitor and evaluate the network governance of actors which daily interact in a defined territorial framework, acting as reliable descriptors of the social processes governing their mechanisms (phase I), and, by consequence, b) how the results achieved are important to understand inter-organizational networks behaviour (phase II). Even if our reflection starts presenting the whole project – both of parts are strictly combined and consequential – in this paper we will explain what we analysed in the second step of the study from a multilevel point of view.

Research first phase has been already described in the cited article (Salvini *et al.*, 2020). Applying *Social Network Analysis* (SNA) strategies and measures on our two case-studies, we understood the network structures of their public and Third Sector nodes, analysing which structural properties are relevant in generating an effective governance (Provan *et al.*, 2005; Wang, 2015; Robins, Bates & Pattison, 2011; Christopoulos, 2008). According to Robins, Bates and Pattison (2011), two endogenous mechanisms have been considered to evaluate the effectiveness of community welfare networks, given particular attention to data related to reciprocity: the *relational embeddedness* and the *structural embeddedness*, as elements contributing to network realization and collaboration, to sharing resources and rules (Granovetter, 1992; Jones, Hesterley & Borgatti, 1997). This part of the study was then aimed at giving a contribution to construct formal procedures, to enhance the network parameters in the evaluation of the *network effectiveness* (Robins, Bates & Pattison, 2011) of systems of relations acting in particular welfare communities. Indeed, the two organizational networks were committed to experimenting a model of coordinated intervention induced by their two corresponding egos – the Caritas centres of Aversa and Benevento. By combining ego-network and whole-network techniques strategies we identified network boundaries, and described the inter-organizational networks functioning.

In this second phase, we refer to and start from the outcomes gained in the first phase. If, as we know, the adoption of forms of distributed governance within inter-organizational networks is fundamental in achieving positive social interventions, the attention is currently focused on how the affiliation of different organizations to specific activities can influence emergent dynamics and behaviours, such as the arising and development of collaborative networks (Provan *et al.*, 2005; Wang *et al.*, 2013; Bodin & Nohrstedt, 2016; McGlashan *et al.*, 2019). We do not only describe the networks properties of the Caritas networks, but we underline the collaboration networks *within* levels and *between* different levels on interaction, deepening a multilevel analysis.

A *Multilevel Exponential Random Graph Model* (*Multilevel ERGMs* or *MERGMs*) is developed here to explore the structural configurations of collaboration networks and community-based tasks in cross-level interactions: this statistical approach has been applied on datasets of 29 and 35 organizations, to show how patterns of actor and task interdependency are predictive factors to influence the formation of collaborative networks, because a good fit among networks nodes appears to be related to a more effective collaboration. A multilevel analysis can help to find out how much the micro- and the meso-level network are intertwined: at a micro-level, network captures the interdependence between the organizations involved (*one-mode network*), at a meso-level, it captures the interdependence between task and actor collaboration (*two-mode or affiliation network*). Finally, we show how cross-interactions and interdependencies among networks and tasks are useful and strategic elements to develop a positive welfare community in a specific territory.

We insert our analysis into a wider discussion stimulated by the application and the effects of the Third Sector Code. The theoretical framework of the whole study is represented by the Third Sector processes acting at local level, in welfare communities, according to the new conditions given by the recent Code precepts. The discussion bases on the consideration that, in Italy, welfare system is built on the wide relationship between public institutions and the detailed activity of voluntary work, acted in a collective fieldwork both at a social, health, cultural, and civil level. The advocacy realized by organizations through their social actions is preeminent in developing social capital.

Scholars wrote about how much informal forms of collaboration, and of interdependence between collective social actors (Raab & Kenis, 2009) – especially within the "solidarity universe" – improve community-building dynamics, enhancing social capital and generating relevant outcomes in terms of sharing resources, support, and knowledge. Recently, Italian Law systematized the role of Third Sector, reorganizing its system, making uniform its diverse entities, and promoting a network approach which has become, in a certain way, more "institutional". Furthermore, the introduction of the Third Sector Code has indicated a central role for networked actions among social organizations, recognizing the main role of collaborations of actors in enhancing civil society, to face current challenges, and to answer to increasingly emerging needs of citizens. This Reform has realized the concept of an *enlarged public sphere*, according to a pattern started with the Reform of V Title of the Italian Constitution, and implemented by the Law 328/2000. Moreover, as in the *Social Origins Theory* theorized by Salamon & Anheier (1998), the link between Third Sector and public sphere has changed into a network of interconnections, along a process that has ended up transforming the role played by non-profit organizations towards public institutions, and therefore the entire welfare community. Despite of the ongoing fragmentation and diversification of regional and local welfare systems, literature demonstrated how the transformations of public sphere (Kazepov, 2009; Wagner, 2000; Bertin & Fazzi, 2010; Salvini *et al.*, 2020) can positively influence the effectiveness of public policies (i.e. *generative welfare*). As Wagner noted in his paper *Reframing "Social Origins Theory": The Structural Transformation of the Public Sphere* (2000), during past years a dynamic of embeddedness of non-profit and public in a common social field developed, increasing the involvement of Third Sector in social intervention activities, in networking, and in providing for the outsourcing of social welfare services which were typical competencies of public institutions. This process created a network of collaboration, defined by a series of economic and social transitions between Third Sector nodes and different levels of government. This is because Third Sector often acts towards several specific needs of citizenship, even before institutions: the outcome is a closer cooperation and interconnection between the two spheres, fostering the inter-organizational network development and the systems of interdependence between collective entities. The fundamental role of Third Sector in communities of welfare has been ratified by the IIV Title of the Code: artt. 55 and 56, stimulating and formalizing practices of co-design and co-planning in social interventions, regulate the relationships between entities and public institutions, particularly through forms of *network governance* intentionally created to achieve social purposes (Raab & Kenis, 2009).

We already assumed that the use of concepts of such as network, network intervention and network governance is fully consolidated, and how they work in the Caritas networks of Aversa and Benevento (Salvini *et al.*, 2020). In this paper we will deepen our study about these networks from a multilevel perspective: we understand (1) why actors choose to collaborate, and particularly why do they choose to collaborate with certain others; we analyze (2) if the condition of affiliation to the same tasks can influence the formation of links between Third Sector organizations at a micro-level; and we test (3) how much the use of ERGMs techniques can be a useful method to understand network governance, network collaboration, and community-building processes. Below our three research hypotheses:

- (H1) in the same local context, organizations which perform the same activities (community-based tasks), are more likely to establish forms of interdependent links;
- (H2) The execution of some specific community tasks is a factor able to influence the structure of the inter-organizational network, through a process of affiliation-based closure;
- (H3) MERGM can represent a feasible analytical strategy to represent multi-level mechanisms of network link formation.

Paper is articulated as following. In Section 2 authors reconstruct the theoretical debate among collaborative networks in welfare communities. The application of the strategies suggested by SNA to analyse network governance is explained, reflecting on the use of ERGMs in understanding inter-organizational networks of community welfare among public, private and Third Sector organizations. The first phase of the study is considered as the starting point to discuss the most recent approaches to inter-organizational networks and their dynamics, such as *affiliation-based closure.* Research hypotheses are presented, and authors introduce data analysis. Section 3 represents the empirical core of the article: setting, participants, data collection, measures, application of ERGMs method for a multilevel analysis, and the final network configuration achieved are detailed and analysed step by step. Results, elaborated with the support of MPNet software, are proposed in Section 4. In Section 5 outcomes are discussed, arguing how a good fit among networks nodes is linked to a more effective sharing (of support, knowledge, resources, activities). In conclusion (Section 6), researchers underline how patterns of actor and task interdependencies are important elements to enhance collaborative networks in a welfare community, and suggest future directions of analysis to networks development and community-building.

2. Theoretical background and research hypotheses

In recent years, some studies have shown that collaborative network governance has a strong normative and cultural appeal, based on the assumption that, practically, an inclusive and horizontally organised network increase the effectiveness of addressing complex and cross-sector problems (Ansell & Gash, 2007; Kickert, Klijn and Koppenjan, 1997; Salvini *et al.*, 2020). For its practical implications, a central challenge for studying welfare community is understanding the mechanisms structuring networks governance, and how they operate and perform supporting societies to respond and identify common solutions to social problems (Provan & Kenis, 2008; Bodin & Nohrstedt, 2016; McGlashan *et al*., 2019).

As mentioned above, in the first part of the work the focus was on two endogenous mechanisms, the *relational embeddedness* and the *structural embeddedness*. The two endogenous mechanisms were identified by Granovetter (1992) and reported by other authors afterwards (Jones, Hesterley & Borgatti, 1997; Robins, Bates and Pattison, 2011). The first one highlights the importance of knowing the strength of the dyadic bonds (measured through *reciprocity*), while the second one focuses on the relevance of the dynamics of composition and interdependence between its various sub-systems (measured through *transitivity*). The data related to reciprocity is particularly important to understand the strength of the ties at the dyadic level (and hence the measure of *relational embeddedness*) and the functioning of the system (Robins, Bates & Pattison, 2011); a low degree of reciprocity indicates a differentiated access to the resources mobilized and a low propensity to the mutual use of those resources. Because of the organizations strong tendency to create ties with their partners of partners, scholars of inter-organizational networks are very interested in dynamics of tie maintenance and formation based on structural embeddedness, defined as closure mechanisms (Robins, Bates & Pattison, 2011; Lomi & Pallotti, 2013, 2013; Bodin & Nohrstedt, 2016; McGlashan et al., 2019). In particular, as Lomi & Pallotti (2013, pp. 202-203) point out, the tendency toward closure mechanism in inter-organizational networks has been interpreted "as consequence of the costs and risks inherent in the formation and maintenance of network ties with partners whose quality, capability, and trustworthiness are only imperfectly observable". To face uncertainty, organizations decide to form new collaboration with their partners of partners based on referrals and shared information (Uzzi, 1996). The results of these studies show that the organizational design of network governance can positively influence these mechanisms and their indicators, generating desirable structural features.

In addition to these two mechanisms, whose empirical results will be presented in the following pages, in order to develop more elaborate accounts of effective collaborative arrangements for complex system, it is important to examine how networks of collaboration among organizations are embedded in, and linked to, other factors in the *whole-of-system*. Specifically, further directions of research should investigate how the distribution of tasks in the system could influence collaborative patterns among them.

Scholars recognize that without understanding the potential interdependencies between inter-organizational networks and the whole system, it is possible to analyse network components only at a single level, but in reality a simplification like this is likely to be insufficient (Bodin & Nohrstedt, 2016; McGlashan *et al.*, 2019).

Advanced model in social network analysis, particularly at the multilevel network modeling (Brennecke & Rank, 2016; Wang *et al.,* 2013; Zappa & Lomi, 2015), are capable to capture in an innovative way the complexity of system interactions. Contractor, Wasserman & Faust (2006, p. 684), as cited in Zappa & Lomi (2016, p. 543) emphasize:

> [O]ne of the key advantages of a network perspective is the ability to collect, collate, and study data at various levels of analysis […]. However, for the purposes of analyses most network data are either transformed to a single level of analysis […] which necessarily loses some of the richness in the data, or are analysed separately at different levels of analysis thus precluding direct comparisons of theoretical influences at different levels.

This new approach to networking therefore considers the model of connection variables and the formulation of hypotheses on variables influence by analysing multilevel network dependencies (Snijders & Bosker, 2012; Wang *et al.*, 2013; Zappa & Lomi, 2016). For this reason, the effectiveness of applying a multilevel network models in diverse fields is strengthening, in order to identify meaningful patterns of interdependencies across multiple levels.

When we talk about *whole-of-system* we mean that collaboration networks are embedded in larger structures that may include multiple levels of action or social system layers (Wang *et al.*, 2013). This multiple level tends to induce dependence relations among participants because, according to Zappa & Lomi (2016) and Feld (1981, p. 1015), "[a]s a consequence of interaction associated with their joint activities, individuals whose activities are organized around the same focus will tend to become interpersonally tied and form a cluster". A specific approach in the analysis of social networks specifies the multiple dependence mechanisms between network ties among actors (in our case-studies, organizations), and the sets of possible *community tasks* (activities, project, program) (Robins, Bates and Pattison, 2011; Bodin & Nohrstedt, 2016; McGlashan *et al.*, 2019; Zappa & Lomi, 2015, 2016). We know that social settings (e.g., a welfare community) are collections of differentiated community tasks (e.g., projects for poor, family, elderly people) in which actors (e.g., public, private and Third Sector organizations) are jointly involved (Mische & Pattison, 2000).

By consequence, in analysing of an inter-organizational setting we can say (hypothesis 1):

H1*: In the same local context, organizations which perform the same activities (community-based tasks), are more likely to establish forms of interdependent links.*

Lomi & Stadtfeld (2014), according to Feld (1981), explain the multilevel analysis introducing a multilevel process of closure by affiliation: *affiliation-based closure*. They defined *affiliation-based closure* as a multilevel network mechanism generating social relations through the execution of the same activities (or tasks) (McPherson, Smith-Lovin & Cook, 2001). Using the SNA vocabulary, we can talk about "closure" because multilevel process induces bipartite clustering, in which the third node of the triad is an object: in our case, it is a task.

The tendency at similarity in the choice and execution of community activities is a factor having a high probability of encouraging the creation of direct ties among the network nodes. This is the reason why a joint affiliation to specific social activities increases the mutual awareness, and provides opportunities for collaborations (Lomi & Stadtfeld, 2014; Wang *et al.*, 2013; Bodin & Nohrstedt, 2016; McGlashan *et al.*, 2019). Talking about clustering in affiliate networks, Hollway & Koskinen (2016, p. 389) confirm this mechanism pointing out that "shared affiliations provide foci or social settings that promote further shared affiliations or even new foci".

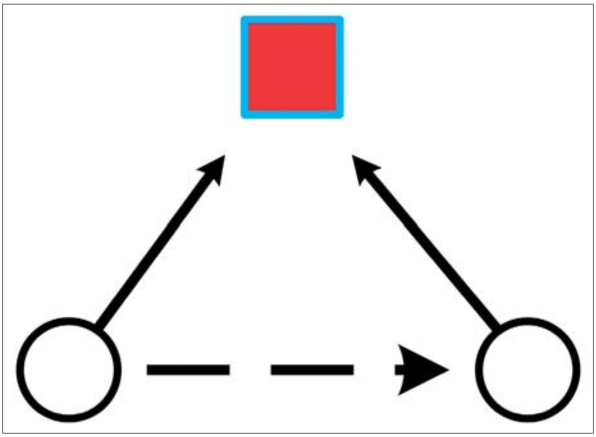
Therefore, we can argue that (hypothesis 2):

H2*: The execution of some specific community tasks is a factor able to influence the structure of the inter-organizational network, through a process of affiliation-based closure.*

The actor-task interaction type refers to the best known forms of bipartite association between rows ("actor") and columns ("tasks") of a two-mode ("affiliation") network. As can be seen in the illustrative image (Fig. 1), the relationships are defined in two different levels of analysis: organizations (represented by the circles) and community level activity (represented by the small square). Sharing the same activities can increase probability that two nodes interact with each other.

In our study there are three levels of analysis: a *micro-level analysis*, which describes the collaboration among organizations, a *meso-level analysis*, which is the activity of an actor into differentiated tasks and *multi-level analysis*, which refers to the interdependence between *micro*- and *meso*-level. The objective of the empirical analysis is to establish whether this particular mechanism of multilevel closure explains the relationship structure of the inter-organizational network, using a *multilevel exponential random graph models* to infer whether actors select certain configurations among other possibilities (Wang *et al.*, 2013).

Fig. 1 – Closure by affiliation. Circles are organizations



Square is a community-based task (adapted from Lomi & Stadtfeld, 2014)

In this paper we introduce *Multilevel Exponential Random Graph* (*Multilevel ERGMs* or *MERGMs*), a class of ERG models for multilevel networks, and we discuss about their importance in understanding inter-organizational network of welfare community among public, private, and Third Sector organizations. MERGMs are useful to investigate the proposed models using simulations and estimation which show that even very simple effects can create network structure both at a single and at a multiple level (Wang *et al.*, 2013; Amati *et al.*, 2019).

Recent application examples of MERGMs concern a considerable variety of phenomena: (i) studies of leadership and the effect of common group factors (e.g., complexity, professionalism and culture) on leadership emergence and performance (Mumford *et al.*, 2008); (ii) collaboration among organization on complex task in disaster management (Bodin & Nohrstedt, 2016); (iii) hierarchical subordination and informal communication relations between top managers in a multiunit industrial group (Zappa & Lomi, 2015); (iv) complex interdependencies between committee collaboration and their actions (McGlashan *et al.*, 2019); (v) bilateral links of States, incorporated into multilateral fisheries agreements, which are in turn grouped around outbreaks represented by similar content (Hollway & Koskinen, 2016); (vi) interplay between formal project memberships and informal advice seeking in knowledge-intensive firm (Brennecke & Rank, 2016).

We can now affirm that (hypothesis 3):

***H3**: MERGM can represent a feasible analytical strategy to represent multi-level mechanisms of network link formation.*

The model interpretation reveals the dependencies among the micro-, meso-, and macro- network level, and provides richer and more detailed descriptions of empirical data. This particular statistical approach has been developed in a complex analytical framework for the Multilevel Networks Analysis (MNA) o for Multilevel Analysis of Network (MAN) (Zappa & Lomi, 2015; Brailly *et al.*, 2016).

3. Data collection, methods, and analysis

3.1. Setting

The research project is about two inter-organizational network of Caritas community: *Caritas of the Diocese of Aversa* (RCA) and *Caritas of the Diocese of Benevento* (RCB). The communities are located in Southern of Italy, in Campania region and are part of Italian Caritas. Both aim to respond to the needs of local communities, they act through multiple activities in partnership with other local organizations, and are supported by a system aimed at charity and altruism.

Recently, Italian Caritas organization developed a further reflection on the importance of improving local networked organizational synergies, in order to respond in a coordinated way to social needs of the communities, beyond the fragmentation and self-referentiality of local welfare agencies. Specifically, the two cited local groups of Caritas have the purpose to generate *networks*, promoting exchange projects and collaborations involving Third Sector entities, local governments, educational institutions and church organizations (Vasca, Riccardo & Capuano, 2016; Vasca *et al.*, 2016). In this generative process, the central nodes of RCA and RCB played a role of animation and coordination, aimed at promoting the creation of networks according to the different project opportunities locally created.

3.2. Network construction and data collection

The first part of the empirical research of the two Caritas networks was articulated in three steps: in the first, the nodes of the two networks under investigation were identified; in the second, data collection operations were carried out; in the third, finally, the data were analysed and interpreted. The first step is to identify the boundaries of the network through the identification of the nodes that make part of it; we proceeded through a combined strategy of ego-network and whole-network. Following the data collection practices in ego-network mode processes, we have identified in the diocesan Caritas of the two territories the *ego node* of the network. This node was asked, using the *name generator* technique, to indicate the alters nodes that are part of its network, through the reconstruction method, indicating the subjects with which this node collaborates for project activities aimed at the communities. The boundaries of the network, therefore, are established by the ego subject, which as such we have defined as "generator node". In this way we built a first-level network, composed by ego and by all the elicited alters at a distance of one link from ego.

At this point, instead of asking ego to indicate the relationship between all alters (as usually happens in the ego network), it was asked directly to the alters by submitting a knowledge questionnaire. They were asked to specify the existence and intensity of the link with each of the alters, choosing one of five increasing levels of the relationship: *no knowledge*, *lack of knowledge*, *good knowledge*, *relationship of exchange*, *cooperation*. The resulting directed graph is what we call "ego-whole" network. In this way, the role of ego was to help define the boundaries of the network, but the relational system emerged only through the interviews carried out with each single alters, according to what is commonly done in whole network surveys; consequently, it was not necessary to use the name interpreter (the tool for the collection of attribute data, i.e. characteristics, of the individual nodes), as is normally done in surveys with ego-centred techniques.

Finally, we asked each node to report the tasks in which they are most involved. They then chose the activities through a list of 12 common *community-based tasks*: (1) seniors care; (2) physically and psychically disabled care; (3) indigents care; (4) prisoners and ex-prisoners care; (5) immigrants care; (6) minors and women care; (7) environment; (8) social and culture promotion; (9) formation and research; (10) arts; (11) enterprise/production; (12) trade. This allowed us to create the matrix of the affiliation network.

Participants were defined as those who are actively engaged in the activity of Caritas in a specific welfare community. At the time of data collection, RCA consisted of 44 nodes, while RCB of 40 nodes. The data gathering was developed through questionnaires to egos and alters during the semester between April 2017 and September 2017. The answers received to the questionnaires correspond to 75.0% (29 nodes) of the RCA nodes and 87.5% (35 nodes) of the RCB nodes.

Table 1 shows the networks composition by type of organization. It is useful to highlight the differences between the networks. Both of them are characterized by a high presence of Third Sector associations and educational and socio-health institutions, with an interesting difference between related to the presence of social cooperatives, which are more consistent in RCB.

Tab. 1 – Composition of RCA and RCB networks by type of organization (category)[2]

		Caritas Network Aversa (RCA)		*Caritas Network Benevento (RCB)*	
Abbr.	*Category*	*Total*	*Answers*	*Total*	*Answers*
AS	Associations	19	11	12	10
CP	Cooperatives	1	1	7	7
OS	Second Level Organizations	0	0	1	1
EM	Religious Organizations	1	1	2	2
GI	Informal Groups	4	4	0	0
ES	Social-health care Institutions	6	6	3	3
SU	Schools-Universities	7	4	6	6
AL	Others (parishes, private organizations)	6	3	9	6
	Total	44	29	40	35

3.3. Descriptive analysis

Descriptive statistics of the Caritas network in the case-studies are in Table 2. We calculated three measures: density network, reciprocity (*relational embeddedness*) and clustering coefficient (*structural embeddedness*).

Tab. 2 – Network descriptive statistics

	RCA	*RCB*
Density	0.58	0.75
Centralization	0.74	0.54
Reciprocity	0.63	0.80
Clustering coefficient	0.65	0.79

The network density is 0.58 per RCA and 0.75 per RCB, considering all types of relationships (no knowledge, lack of knowledge, good knowledge, relationship of exchange, cooperation). Generally, it can be said that, especially for RCB, the density assumes rather high values; obviously the centralization index is higher where the density is lower, especially in RCA, where therefore the relationships depend more on the activity of a limited

[2] They are distinguished by their ego networks generated by the Egos of RCA and RCB (total), and by the number of nodes responded to the questionnaires (answers).

number of nodes. Density and centralization also change considerably if one considers the values related to the different types of relationships: the former decreases significantly if one considers relationships based on exchange, collaboration and good knowledge, while the latter increases in correspondence with the decreasing density.

From the analysis of these data it is possible to highlight that the two networks considered have characteristics, albeit in different intensities, which confirm what we already know about the networks mobilised by this type of organization (Robins, Bates & Pattison, 2011). Indeed, we are in the presence of generally high levels of mutual knowledge, but also of low levels of collaboration and exchange, which are normally polarized around the activities of a few organizations – generally those more structured and more consolidated from the point of view of human and economic resources management.

The level of *relational embeddedness* in both of the networks has been analysed through the degree of *reciprocity* (Robins, Bates & Pattison, 2011). Reciprocity is important to understand the strength of ties at the dyadic level and the functioning of the system: as already said, a low degree of reciprocity indicates a differentiated access to resources, a low inclination to their mutual use, and it is an index of low agreement about the contents, the objectives and the rules designing the relational structure of the whole network (Robins, Bates & Pattison, 2011). If we consider all the links, even those based on simple knowledge, reciprocity is rather high, a sign that a general, albeit generic, mutual knowledge of organizations in the territory is consolidated.

The combined analysis of the parameters shows that, in general, the two networks (especially for RCB) have a good level of reciprocity which, however, decreases significantly when only those relationship sets based on good knowledge and exchange of information and resources are considered. This does not mean that there are no flows of resources in the network, but that these flows do not involve reciprocal mechanisms. Moreover, the lack of reciprocity of relations is considered an indicator of a low level of sharing and agreement on the contents and objectives that govern the existence of links, as well as on the rules that should inform the whole system of relations (Robins, Bates & Pattison, 2011).

The *clustering coefficient* measure analyses the *structural embeddedness*. With these prerequisites, we should not expect particularly positive performances with reference to the clustering coefficient in reality, it is possible to observe that the two networks have a very high global coefficient, which remains consistent in the systems of good knowledge and collaboration relations – while it is reduced, as expected, in the exchange relations. The high level of clustering coefficient suggests that the structure of the two networks is sensitive to the formation of "clouds" of nodes, based on aggregation criteria. This means that the potential for mobilization – and cohesion – of networks is high. At the global structure level, the presence of a high clustering could favour mechanisms of "closure" of open triads, and therefore the achievement of higher levels of cohesion of the network. This eventuality certainly appears to be desirable if the network governance concerns projects aimed at achieving specific goals, and relations systems are empowered by formal agreements (Provan & Kenis, 2008).

3.4. Analytical Methods

In the second phase of the study we analyse our datasets applying recently developed *Exponential Random Graph Models* (ERGMs) for multilevel analysis.

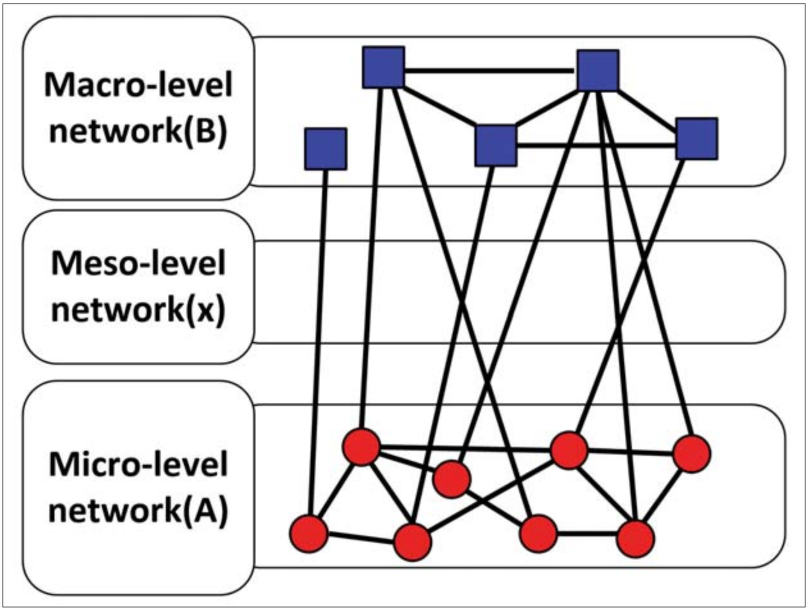
Introduced by Frank & Strauss (1986), Wasserman & Pattison (1996), and originally developed to examine single-level networks, ERGMs represent a family of probabilistic statistical models for the analysis of complete social networks, not longitudinal, through which an exponential probability distribution of the graph is expressed (Robins *et al.*, 2007).

ERGMs are useful to obtain insight into underlying processes that create and sustain the network-based social system (Lusher, Koskinen & Robins, 2013) and assumes the interdependence between the dyads. The model expresses the probability that a particular configuration of the network will occur (e.g. existence of a link) as a function of a certain number of parameters that can be represented by relational variables (structural or network effects such as *reciprocity*, *density*, *transitivity*, *cyclicality*) and attribute variables (*gender*, *age*, etc.) and whose estimation can be obtained through specific software.

Exponential random graph models have been extended as MERGMs to assess network interdependencies across multiple level of network data (Wang *et al.*, 2013). Multilevel models specify a set of *multi-level* configuration, that are expected to explain micro-level of outcome variables. MERGMs add the possibility of testing hypothesis about how the presence of ties (i.e., inter-organizational network) among lower-level actors or units (e.g., organizations) depend on the ties among higher-level actors or units (e.g., among organization and tasks) (Wang *et al.*, 2013).

The outcome tie variable predictors are local configurations of network ties. Local configurations represent relational mechanisms – like reciprocity and transitivity, previously mentioned – and are defined across multiple levels. Parameters correspond to specific relational mechanisms taking place at *within* as well as *across* network levels. By including micro-level as well as multi-level structural parameters, we estimate their contribution as drivers of local structural patterns (Robins *et al.*, 2007; Brennecke & Rank, 2016). For a two-level network, we have the micro-level network as *network A*, the macro-level network as *network B*, and the meso-level bipartite network as *network X*. This flexible data structure, as in Figure 2, generalizes the multilevel model.

Fig. 2 – Multilevel network schematic diagram



Source: adapted from Wang *et al.* (2013)

In our research, the interdependence assumptions considered concern exclusively the micro- (A) and multi-level (A&X). Multilevel analysis starts with the identification of a set of micro-level actors (organizations), of an affiliation network (actors and community-based tasks), and of network ties within and between elements of settings. Supposing A as the inter-organizational network, and X as the set of community-based tasks, A&X stands for relations between actors and tasks across the different levels; there is a cross-level relation affiliating organizations to specific tasks that associate elements of A and elements of X (Zappa & Lomi, 2015).

Using typical *MERGMs* nomenclature to represent realizations of our network variables, we have $\{A, X\} = \{a, x\}$; and $(Yi = yi)$ represents the attribute variables for organizations in collaboration network (*A*). The MERGMs for our networks can be expressed as:

$$P_\theta(A = a, X = x \mid Y = y) = \frac{1}{k} exp \sum_Q \theta_Q z_Q \ (A,\ X,\ Y)$$

We model here the structure of collaboration network among organizations (*A*) and choices of community-based tasks (*X*), treating attributes (*Y*) as exogenous:

The Equation represents the probability distribution of all possible graphs on the same set of nodes. The probability of obtaining a specific graph (i.e., that A and X – is equal to the sample distribution *a* and *x*) depends on the presence of various configurations included in the model.

– *Q* indicates network configurations of type Q – as discussed further – within which all tie variables and nodal attributes are considered as conditionally dependent following the hierarchy of tie dependence assumptions, that is, the occurrence of a tie is dependent on the existence of other network ties;
– z_Q (*A*, *X*, *Y*) are network statistics counting the number of configurations of type Q. The statistics count, for each actor in the network, the number of configurations of each type in which the actor is involved, for example, the number of reciprocal ties including actors;
– θ_Q are parameters associated with configurations of type Q. A positive and significant estimate of θ_Q suggest the corresponding configurations happens more than one would expect from random given the rest of the model, while negative estimates mean the opposite. A parameter is considered significant if its estimate is more than twice the estimated standard error;
– κ is a normalizing constant summing over the entire graph space bounded by the number of nodes in the network.

As several authors note (Snijders *et al.*, 2006), MPNet software was used to conduct analyses which implement the MCMC maximum-likelihood estimation algorithm for ERGMs (Wang *et al.*, 2009, 2013). We assess the goodness of fit of the final model, by comparing the full suite of non-estimated statistics available in MPNet of the estimated model was assessed by simulating a high number of graphs from the fitted models. We compared the characteristics of the simulated graphs to the characteristics of the observed networks.

3.5. Network configuration

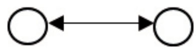
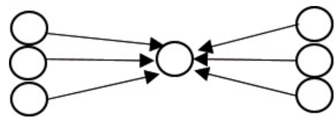
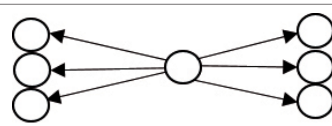
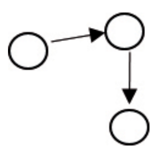
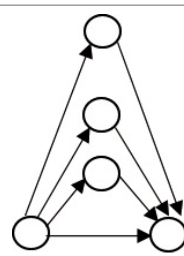
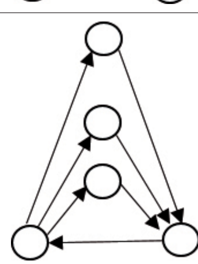
Table 3 provides a list of selected MERGM configurations or parameters we used in our models with their visualization and potential interpretations. The network configurations are local network patterns which provide explanations about the reason why ties might be present in a network, how ties might form particular local network patterns, and how ties might be associated with actor attributes.

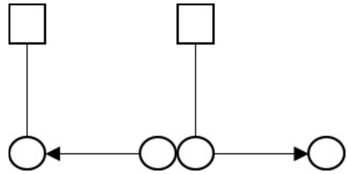
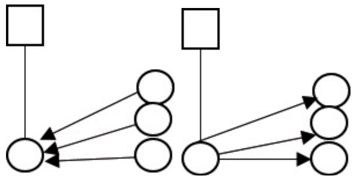
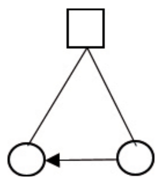
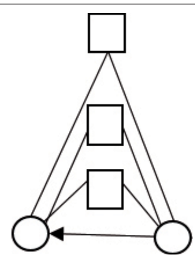
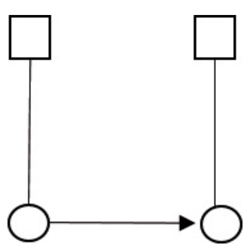
Concerning the level parameters selection capturing the network structure, we accord to model specifications used in several studies on inter-organizational network (e.g., Lomi & Pallotti, 2013; Bodin & Nohrstedt, 2016; McGlashan *et al.*, 2019; Brennecke & Rank, 2016; Amati *et al.*, 2019). Since our research is mainly concerned with the multilevel interdependencies characterizing actors and task network, micro-level configurations are included to capture in a stochastic model the mechanisms of relational and structural embeddedness. In our study, attribute variables were not included because the model was already quite complex and with satisfactory adaptation values.

Thus, the model includes specific configurations of two levels:

– *micro-level configurations* investigate relations within the one mode network A. The simplest effects are *arc* and *reciprocity*. The former represents the tendency of organizations to establish ties with other members of the network; the latter the tendency to reciprocate existing incoming ties. This is often positive, suggesting that reciprocated ties are very likely to be observed for positive affect networks;
– *popularity* mechanism describes tendency of inter-organizational network members to be popular, i.e. to be considered as connection partners by many organizations. A negative popularity spread parameter indicates that most actors have similar levels of popularity (the network is not centralized on in-degree). *Activity* mechanism represents the propensity of being active, i.e. to connect with many organizations; a negative activity spread parameter indicates that most actors have similar levels of activity (the network is not centralized on out-degrees). *Two-paths* mechanism indicates tendency to connect with other organizations, and to be sought as connection partners by other organizations. If this parameter is significantly positive, it can suggest that the most popular actors are the most active ones (i.e., a positive correlation between incoming and outgoing distributions);

Tab. 3 – Selected MERGM configurations and possible interpretations

Micro-level effect (A)		
Effect (MPNet command)	*Configuration*	*Interpretation*
Arc		Tendency of inter-organizational network members to connect with each other
Reciprocity		Tendency of members of the organization's inter-organizational network to connect with colleagues, reciprocally
Popularity Spread (AinS)		Tendency of inter-organizational network members to be popular, i.e. to be sought as connection partners by many organizations
Activity Spread (AoutS)		Tendency of inter-organizational network members to be active, i.e. to connect with many organizations
Two-path		Basic tendency of inter-organizational network members to connect with other organizations and to be sought as connection partners by other organizations
Transitive closure		Tendency of inter-organizational network members to connect with "partner of partners"
Cyclic closure		Tendency of organizational network members to connect with other organizations in small groups without any expectation of being reciprocated

Multilevel effect (A&X)		
Affiliation based popularity of activity effect (In-2StarAX, Out2StarAX)		Tendency of actors who are linked to a task to establish in- and out- connections with other actors
Affiliation based popularity of activity effect (AXS1Ain AXS1Ain)		
Affiliation based closure or homophily by a common affiliation (TXAXarc)		Tendency to connect with organizations carrying out one of the same activities. The dependence is that of Markov's models, with the difference that the third node is not an actor in network A, but a task)
Affiliation based closure, or homophily by a common affiliation (multi-level social circuits) (ATXAXarc)		Tendency for nodes sharing multiple affiliations to connect to each other, or homophily by multiple common affiliations
Assortativity based on popularity in affiliation network or precondition for cross-level closure (L3XAXarc)		Tendency for two similar nodes (compared to the number of tasks performed) to connect to each other. Indicates that the probability that two nodes with the same number of tasks performed are more likely to connect than two nodes with different numbers of tasks do

Note: circles represent organization of Caritas Network, and squares represent community-based tasks.

Source: Wang *et al.* (2013)

– more complex effects involve more than two nodes: *transitive closure* and *cyclic closure*. Transitive closure models the tendency to connect with partner of partners. A positive effect here indicates that in data there are a high degree of closure, or multiple clusters of triangles; whereas cyclic closure tells us that members to connect with other organizations in small groups without any expectation of being reciprocated. A negative effect indicates tendencies against cyclic triads. This closure mechanism is based on the concept of generalized exchange, a triadic pattern of network ties that does not involve immediate needs for reciprocation. Generalized exchange defines a situation in which "the recipient of benefit does not return benefit directly to the giver, but to another actor in the social circle. The giver eventually receives some benefit in return, but from a different actor" (Molm, Collet & Schaefer, 2007, pp. 207-208). It plays an important coordination role in inter-organizational networks because "any particular exchange or transfer of resources occurs in a context of other exchanges and transfers" (Lazega & Pattison, 1999, p. 68). A negative estimate in conjunction with positive triangulation estimates indicates that 2-paths tend to be closed (i.e., triangles are formed);
– *multilevel configurations* of cross-level network A&X "express tendencies for structural [configurations] to be associated across both levels simultaneously" (Wang *et al.*, 2013, p. 99). The set of multilevel effects describe how the internal tasks of the organizations (ties in the two-mode network) affects inter-organizational relations (ties in the one-mode network). *Affiliation based popularity of activity effect* (In2StarAX and Out-2StarAX) is the tendency of actors who are linked to a task to establish in- and out- connections with other actors. *Affiliation based closure or homophily by a common affiliation* (*TXAXarc* and *ATXAXarc*) captures the tendency of network organizations carrying out the same project activities to connect with each other. Specifically, the parameters indicate the likelihood of ties to occur between organization with same activity more or less often than expected by chance. The dependence is that of Markov's models, with the difference that the third node is not an actor in network A, but a task (Lomi & Stadtfeld, 2014; Amati *et al.*, 2019). This latter mechanism is the effects of central interest in the analysis, as it explains the mechanism we have introduced, the *affiliation-based closure*. Finally, *assortativity based on popularity in affiliation network or precondition for cross-level closure* (*L3XAXarc*) describes how actors that are connected the same number of tasks (in the two-mode network), tend to connect more easily.

4. Results

The parameter estimates for MERGMs of RCA and RCB are shown in Table 4, where there are list the effect names, parameter estimates, estimated standard errors. For the convergence of the model, t-ratio must be less than 0.1. A parameter is considered significant and marked by "*", if the ratio between the estimates and estimated standard errors is greater than 2.0 in absolute value. We also included non-significant effect in order to provide an adequate fit to the overall network structure. Results in Table 4 are arranged in the part representing micro-level structural effects (inter-organizational network A), and in the second part representing multilevel structural (inter-organizational network and affiliation A&X) contained in the models to test our hypothesis.

Tab. 4 – MERGM parameter estimates

		RCA		*RCB*	
	Effects	*Est.*	*Std. Err.*	*Est.*	*Std. Err.*
Inter-organizational network (A)	ArcA	-1.717	0.744*	-2.493	0.318*
	ReciprocityA	1.266	0.31*	1.647	0.246*
	In2StarA	0.068	0.036	0.065	0.021*
	Out2StarA	0.053	0.036	0.136	0.017*
	TwoPathA	-0.075	0.018*	-0.056	0.013*
	Transitive-TriadA	0.073	0.027*	0.001	0.016
	Cyclic-TriadA	-0.076	0.03*	0.008	0.03
Inter-organizational network and affiliation (A&X)	In2StarAX	-0.026	0.169	0.043	0.066
	Out2StarAX	0.170	0.18	0.050	0.074
	AXS1Ain	0.096	0.075	-0.015	0.033
	AXS1Aout	-0.031	0.084	0.033	0.034
	TXAXarc	0.103	0.033*	0.125	0.022*
	ATXAXarc	0.072	0.078	-0.057	0.058
	L3XAX	-0.047	0.019*	-0.032	0.009*

A separate text file has to be created for each network. The file consists in a square binary matrix for network A and a rectangular binary matrix for network *X*. For instance, the matrices used in our empirical cases are as follows:

– A is a square matrix of size 29x29 (for RCA) and size 35x35 (for RCB). Its generic entry has value "1" if there are exchange and collaboration ties, and value "0" for the lack of knowledge. The diagonal values are set at 0 by default;

– X is a rectangular matrix of size 29x12 (for RCA) and size 35x12 (for RCB). Its generic entry has value "1" if the row elements (organizations) are affiliated – i.e. particularly engaged – to the column element (tasks) and 0 otherwise. Each organization can be engaged in many activities according to different intensities (or levels of commitment)[3], but to simplify we made the binary matrix giving value mainly to activities where they are engaged in with high intensity (*a lot of commitment and total commitment*).

Findings are now described accordingly to the two levels of analysis.

Inter-organizational network A. The effects capturing the structure of *inter-organizational* networks are included as control variables, some of which confirm the results of the previous study on relational and structural embeddedness mechanism (Salvini *et al.*, 2020). In both of the networks, we find tendency toward *Arc parameter* negative and significant, typical feature of empirical networks indicating the existence of a ceiling effect to establishing inter-organizational ties (Zappa & Lomi, 2016). So there is no general drive to establish forms of collaboration between the various organizations of the two Caritas networks.

Data, however, have strong evidence for *reciprocity*, showing that there is a high probability to be confirmed by organizations ties (i.e. are reciprocal). This is a very standard finding in inter-organizational networks, where the most important relational behaviours, characterised by collaboration and partnership between this type of organization, require mutual knowledge (Amati *et al.*, 2019; Lomi & Pallotti, 2013; Robins, Bates & Pattison, 2011). The qualitative implication of good level of reciprocity is that between organizations committed to satisfy the needs of local communities can lead to the sharing of knowledge and resources in general, and therefore in tackling together needs and difficulties related to shortage of material and immaterial resources typical of this set of organised realities (Salvini, 2011).

The effects related to incoming and outgoing degree (*In2starA* and *Out2starA*) indicate the convergence of links to a small number of organizations in the two collaboration networks. Evidence of centralization tendencies based on in- and out- degree is also found. It is possible to observe that the *popularity* parameter (*In2starA*) is positive for RCA and RCB, although it shows levels of significance only for RCB. This means that both networks are characterized by high levels of centralization, but in RCB, more than in RCA, there is a generalized trend spread only for several nodes. The *activity*

[3] Levels of commitment: *no commitment*, *little commitment*, *sufficient commitment*, *a lot of commitment*, *total commitment*.

effect (*Out2starA*) is positive and not significant for RCA, and positive and significant for RCB: this indicates the presence in RCA of few organizations that value and encourage forms of collaboration with others, while in RCB network there is a greater trend towards activity and hence some organizations are more active in seeking collaborative relationships.

The tendency of the most popular actors to be also the most active ones in searching for collaborations with other nodes of the network, is shown by the *TwoPathA* parameter. In this regard, it is possible to note that, both in RCA and RCB, this value is negative and significant, demonstrating that there is no positive correlation between ingoing and outgoing degree. The negative and significant correlation between the internal and external degree in RCA and RCB networks means that members who have received more collaborative links are less likely to send them. This could represent different levels of influence or experience in the networks, so that the more established and stronger members are more popular recipients of collaborative applications, but less active and perhaps more selective in sending and actively forming collaborations with others. This, of course, is not a valid consideration if we look at the Caritas ego node as it will have the same number of incoming and outgoing links.

The positive *Transitive-TriadA* property in both networks is indicative of two multiple paths leading to the formation of a direct link. This is an efficient collaborative structure, as it indicates a "shortening" model of collaboration between different organizations in the network. For example, if organization i creates a collaborative link with j, which already has a collaboration with y, the initial node *i* could eventually create collaborations with *y*. While both community networks have shown patterns of transitive collaboration, albeit significantly only for RCA, the same network has coupled this with a significant negative cyclical pattern.

Now, observing together the transitivity effects of the network (*Transitive-TriadA* and *Cyclic-TriadA*) in Table 4, it is necessary to highlight a certain divergence in networks data. In RCA network, the value of transitive triad is positive and statistically significant, while the cyclic triad is negative and statistically significant. Together, the two parameters provide clear indications on the triangulation effect of this network: they confirm the high level of clustering (i.e. the propensity to generate collaborative groups between organizations), but the negative value of cyclic triangulation tells us that this level of clustering is characterized by a certain local hierarchy, with a high level of brokerage (i.e. only some nodes determine the connections). There is a low level of generalized exchange, since the flow of information and resources is allowed by few actors in the network. In RCB, the two effects are both positive and not significant, which means that the closure effects cannot be explained through these two mechanisms, but rather through simple mechanisms related to degree, centralization and reciprocity (Lusher, Koskinen & Robins, 2013).

Inter-organizational network and affiliation A&X. Estimates associated with multilevel mechanisms are reported in the lower part of Table 4. The interaction effects between the two inter-organizational networks and the affiliation network include values of affiliation based closure (*AXS1Ain* and *AXS1Aout*) or homophily by a common affiliation (*AXS1Ain* and *AXS1Aout*). Note that none of the parameters are significant. Removing it from the model leads to a poor fit on the internal layer distribution in degrees of the inter-organizational network. Having the homophily affiliation within the statistics explicitly shapes organizations with multiple community tasks, which are the organizations that carry out more activities in the territories.

The positive estimate of *AXS1Ain* suggests that the largest organizations in terms of the number of activities carried out, tend to receive more nominations as collaborators from other organizations. However, on the contrary, as indicated by the negative parameter *AXS1Aout* tells us that they tend to nominate fewer collaborating organizations. Observing the non-significance of these parameters for both networks, we can argue that there is no particular evidence to show how the execution of one or more tasks (i.e. the links in the two mode network) generates collaborative links between the organizations of the two Caritas networks. This interpretation has an empirical and theoretical sense, as organizations carrying out many different activities may be the dominant nodes of the network with respect to their history, good reputation and having more material and immaterial resources, but they themselves have not felt the same urgency to collaborate with others.

Concerning multilevel affiliation, we notice that *TXAXarc* (*affiliation based closure or homophily by a common affiliation*) and L3XAX (*assortativity based on popularity in affiliation network*) are, in both network, statistically relevant. These configurations capture the first two hypotheses (H1 and H2). Positive and significant value of TXAXarc indicates that organizations of Caritas network holding similar community-based tasks display a tendency to develop collaborative network. This value, from a qualitative point of view, tells us that the creation of links can be attributed to the sharing of the same community tasks. For example, if *i* is an organization that deals with minors it will tend to collaborate more with y if the latter deals with minors, compared to *j* that deals with immigrants. This data, in the context of network governance (Provan & Kenis, 2008; Robins, Bates & Pattison, 2011; Christopoulos, 2008) could be of fundamental importance as the flows of knowledge and resources could guarantee an efficient system, as it is based on collaboration between similar associations and therefore with the same targets, need, resources and goals. Also with the risk of compartmentalized of interventions, thus fragmenting the complex network and not allowing the connection between different sectors. In this case the role played by brokerage organizations – that act as a bridge to the networks – is important, connecting different areas of intervention.

The multilevel closure suggests a clear evidence of social selection process as closure mechanism in our data. However, a not significant value of *ATXAXarc* effect tells us that this closure value stops at a certain point. Finally, as we said, the *L3XAX* configuration refers to tendency for two similar nodes (compared to the number of tasks performed) to connect to each other (assortativity). The negative and significant estimation of the parameter for both network suggests that there is a dissortativity effect: organizations that carry out a large number of activities prefer to connect to organizations with a low number of activities and vice-versa.

5. Discussion

The study about two Third Sector inter-organizational networks in Southern Italy – the Caritas networks of Aversa (RCA) and Benevento (RCB) – combines the use of the advanced instruments of Social Network Analysis perspective with the application of the most recent changes in Italian volunteering: it shows how a multilevel analysis be a useful trick in understanding inter-organizational networks' structure, and the social (and systemic) processes underlying a community configuration. The reflection refers indeed to theoretical framework about voluntary work given by the actualization of the *enlargement of the public sphere*, as ruled in Italy by the Third Sector Code in network governance, through the adoption of forms of distributed governance within inter-organizational systems.

RCA and RCB have been firstly understood through the strategies of SNA, assuming that this perspective represents an adequate tool to investigate the way how the structures of links can influence network governance. Network parameters were used to evaluate the network effectiveness of systems of relations between nodes (organizations, entities) operating in a determined welfare community, to strategically identify structural properties having a higher performance (an ability) to activate effective governance (Provan & Kenis, 2008; Salvini *et al.*, 2020). Secondly, as presented in this paper, researchers develop results previous achieved, applying a multilevel analysis to investigate the cross-level links between organizations and community-based tasks. Structural configurations in cross-level interactions are based on a constraint dependency approach, for which network constraints are interdependent not only *within* levels but also *between* levels. We introduce the methodological contribution of MERGMs (Wang *et al.*, 2013; Bodin & Nohrstedt, 2016; McGlashan *et al.*, 2019): we analyse both the micro-level of RCA and RCB, and the meso-level linking the organizations to community-based tasks, in order to understand if affiliation network is able (or not) to impact the diffusion of links between different network levels. And, by consequence, if interaction levels may influence the whole system of relations among Third Sector organizations, enhancing the empowerment of collaborative network in a welfare community. Indeed, operating specific community-based tasks and the need to find different forms of collaboration among actors are primary attributes of Third Sector systems, which determine informal relations and exchanges among typical nodes of voluntary work.

Literature underlined the great relevance of networking dynamics for Third Sector and community empowerment. Structured social relations (networks) are fundamental instrument for explaining the way how society (or a community) works – its level of social cohesion, governance, and social capital – even more than studying personal characteristics of its members (Wellman & Berkovitz, 1988). The importance of networking is daily assumed in many fields, not only in the economic area but also in social intervention: in volunteerism, more or less formalized forms of interactions among organizations see the collaboration of professional entities (or individual actors), to offer services and support to the more increasing complexity of social needs, and to gain a more effectiveness in organizational activities. The link between social capital of a specific territory, and the ability of actors to establish relationships, creating structures of interdependence and collaboration (inter-organizational networks), promotes network solidarity (Psaroudakis, 2011; Salvini & Psaroudakis, 2016). Networking is therefore as a strategic asset to improve not only Third Sector development, but community-building too. Moreover, in this field networking processes has been highly stimulated by the recent Third Sector Code, which has "institutionalized" the usefulness of network activities (i.g. partnerships, agreements, collaborations, shared initiatives), through practices of co-planning and co-design in social interventions.

Nevertheless, from an analytic point of view, there is a lack in inter-organizational studies about a multi-level comprehension of Third Sector networks and their direct and indirect impact on community functioning: only a few of empirical research gave it a strong application (i.e. McGlashan *et al.*, 2019).

We recognize here their importance, since (Third Sector, but not only) organizations are complex systems embedded in such particular frames of activities, whose interconnections can influence also the structure of external networks. This is the reason why we propose an integrated analytical framework to explicitly measure multilevel network dependencies: technically, model used shows a) the mechanisms (*relational* and *structure embeddedness*) at the micro-level, and b) how much the meso-level affects the micro-level structure (*affiliation-based closure mechanism*). Data collected provide a documentary evidence of the empirical value of this method: results validate general findings achieved during the previous phase of the study, and the solidity of ERGMs techniques to understand network behaviour.

The two levels of data analysis confirm our three research hypotheses, and suggest that considering organizations as two-mode or multilevel systems better works to study Third Sector networks. Outcomes presented are strictly connected, as discussed below.

We start from networks description and configuration. RCA and RCB have similar features in identity, mission, activities, field of social intervention, and subjects. Both of them are informal Third Sector networks, voluntarily operating according to principles of solidarity, gratuitousness, attention to people needs – especially citizens experiencing conditions of particular weakness. Interpreting their SNA parameters, they work and collaborate in a similar way (as a "core-periphery" or a "hub and spoke" network): if hubs emerge from interactions, differences come from their own configurations (and nodes) having any impact on networks functioning.

Their collaboration is due to organizational role of some specific entities, or to the coordinated and reciprocal action of several organizations. A high clustering level (as in RCA and RCB) is certainly useful to achieve forms of governance, as long as centralization degree is not too high. RCB has attributes more consistent with a "hub and spoke" functioning of governance, while RCA structure is even more dependent on the activity of a few organizations. In situations in which networks are mainly informal and gradually activated accordingly to specific project opportunities, scholars (particularly Krebs & Holley, 2006) argue that it is more desirable to maintain a "hub and spoke" or "core-periphery" network model, because a hierarchical configuration could guarantee a network solidity, based on the focused and interconnected activity in its core. The core may benefit from the "targeted" collaboration of peripheral nodes to particular activities, providing new resources and innovative tasks, and from the mobilization capacity distribution among other organizations. Entities having less centrality have to be reinforced, gradually assuming roles of coordination and activation: fostering the realization of multi-hub configurations, it is possible to enhance networks characterized by complementarity (Krebs & Holley, 2006; Anklam, 2007).

We come now into focus of discussion.

Hypothesis 1. The properties of RCA and RCB networks are useful to better explain why actors choose to collaborate, and particularly why do they choose to collaborate with certain others. Multilevel analysis is helpful in understanding if there is a higher possibility for network nodes to collaborate with organizations having the same community-based tasks: within a shared local context, organizations promoting the same activities based on similar advocacy, are more likely to establish forms of interdependent relationships.

Effects capturing the inter-organizational network structure at a micro-level (level effect A) are included in our analysis as control variables, confirming results gained during the first phase of the study (Salvini *et al.*, 2020), and the rules underlying the structure of collaboration networks between volunteer organizations. We find negative tendencies of *arc* parameter, positive tendencies of *reciprocity* parameter, and negative tendencies in *TwoPathA* of both RCA and RCB. Nevertheless, as indicated by parameters of popularity (*In2StarA*) and activity (*Out2StarA*) – positive for both of networks, even if relevant only for RCB – there is a tendency to centralization: some nodes seem to be particularly attractive, and active in searching for links. We can affirm that RCA and RCB are centralized networks, and the role of hub played by ego-nodes makes the whole system particularly dynamic. Moreover, transitivity measures (*Transitive-triadA* e *Cyclic-triadA*) offer further elements to evaluate our case-studies. Networks data are different: in RCA network building is strongly influenced by a local transitivity process (even if it shows a negative and relevant *Cyclic-triadA*), whereas in RCB there are no significant tendencies about transitivity (even if values are good). Connecting micro- and meso-level (A&X), the interdependency is demonstrated by TXAXarc parameter. According to hypotheses 1 and 2 (see below), a positive and relevant TXAXarc value shows the inclination of similar organizations to know, and to collaborate with each other.

An interesting info is given by a negative and significant *L2XAX* value (*assortativity*), because smaller organizations (carrying out a low number of activities) tend to connect with the bigger ones, and vice-versa. This is important in community management, because it explains how smaller entities need (and look for) forms of support by consolidated actors, which have more experience and resources. Instead, big organizations are not interested in connecting to nodes with a similar range of activities, but are more willing to establish links with emergent or sector-based realities.

Hypothesis 2. We analyse if the condition of affiliation-based closure mechanisms can influence the formation of links between Third Sector organizations at micro-level, and we discuss how network initiatives of two Caritas nodes are aimed at intensifying connections.

The positive value of *TXAXarc* used in hypothesis 1 validates hypothesis 2 too. An *affiliation-based closure* mechanism (homophily by a common affiliation) acts, and organizations dedicated to specific activities seem to be more linked (in a reciprocal way): an engagement activity criterion is capable of producing a social selection process among network organizations. There is indeed a wider and more effective collaboration based on particular community-based tasks and a sharing of resources (of support, knowledge, activities, practices) among similar organizations, having good effects on welfare community processes. According to literature (Zappa & Lomi, 2016), this data confirms the role of social intervention (*community tasks*) as social foci: having the same social purpose based on the sharing of activities may increase the adoption of general exchange practices in an organizational framework. Affiliation dynamics, homophily, and nodes similarity can improve community structure and system efficiency, but at the same time there are risks of fragmentation among volunteer organizations, and of polarization of interventions to specific and limited sectors. We recognize the role of brokers played by some actors, connecting different activities and nodes, and the importance of fostering the sense of collective identity between community members (public, private, and Third Sector entities) to develop a network culture. Nevertheless, a similarity in identity may be an important reason to collaborate: even if their network configurations are different, RCA and RCB have a sort of "collective" identity because of their common belonging to the same territory, and to a shared ethical framework. Expressing their advocacy and mission, initiatives and collaborative experiences seem to be acted in an equal way.

Thanks to our analysis, we have more arguments to affirm that the further development for the two inter-organizational Caritas networks should be the growth and maturation of their networks relations, to enhance reciprocity in exchanges and forms of collaboration.

Hypothesis 3. Finally, we test how much the use of ERGMs techniques can be a useful method to understand network governance, network collaboration, and community-building processes.

This study introduces an innovative methodological contribution in the use of *Multilevel ERGMs* in Third Sector and in community welfare studies: literature about inter-organizational networks operating in welfare contexts is not so vast, especially referred to Italian territories, and there are only a few applications of Multilevel Exponential Random Graph Models in Third Sector networks oriented to solidarity and community-building. We use MERGMs methods to represent multilevel mechanisms of network tie formation: they are helpful in understanding that links emerging from different organizational levels interact in a complex way. By this, strategies usually adopted to analyse inter-organizational networks can be included in a different conceptual framework, allowing a more accurate interpretation of the sense of empirical SNA parameters in Third Sector field. Limits can be about a deeper understanding of the dynamics of activities selection able at creating collaborative links, and of the nature of relationships related to these mechanisms. For this reason, in conclusions we suggest the application of a mixed-methods approach, because the typical instruments of qualitative methods – added to SNA measures and techniques – can be helpful in investigating the importance of influence processes among network organizations.

6. Conclusions

At the end of the study we validated our three research hypotheses and demonstrated that interconnections among Third Sector entities and community-based tasks interdependency are relevant to improve collaborative networks in a welfare community.

Literature explains that promoting collaborative networks generates positive outcomes for community-building. According to Anklam (2007), a networking activity creates concrete, innovative and multilevel benefits, expressed both at the human capital, the structural capital, and the relational capital level. But even if recent Law promotes the assumption of network strategies, resources are limited and organizations must obtain even more professional skills to be competitive, and act in a specialized way, generating fragmentation: this is the reason why we assist to an increasing isomorphism among Third Sector entities and economic actors, based on a similar identity, and on the meaning of a so-called "new" volunteerism. By consequence, it is necessary to enhance practices of network development in Third Sector organizations, to create solid and efficient networks and to foster the spread of information and resources. Investing in creating collaborative relationships based on trust (reciprocity, transitivity) can reduce this sectoral approach, and improve community welfare. Particularly, in hierarchical networks as RCA and RCB are, it is not possible taking for grant that sharing the same activities (community-based tasks and mutual affiliation) is the only element aimed at generating a collective network identity, a network culture, or a network effectiveness related to coordination practices.

Research presented in this paper offers important implications in Third Sector governance. In conclusion we suggest the opportunity of increasing the use of an advanced SNA perspective combined with the tools of non-standard methodology, in order to identify positive dynamics and critical issues for collaborative networks development in Third Sector fields. We demonstrated that information collected and analysed through ERGMs techniques are particularly interesting in understanding the development dynamics of communities. A mixed-methods approach (Bellotti, 2015) can help policy makers to consider the organizational structure of networks, their nodes and relations, and which kind of governance may be the most efficient one to improve network strategies for community-building. Results gained through the use of SNA parameters and techniques – system description, and its conditions of subsistence and development, also in a longitudinal way – can be enriched by applying the typical instruments of qualitative methods such as interviews and focus group. We refer here to the theoretical and methodological perspective of *Symbolic Interactionism* and *Grounded Theory* (Blumer, 1969; Charmaz, 2008): if every action is a relationship, and is embedded in a collective dimension (the community), a mixed-method analysis will be focus also on the "meanings" of networking.

Therefore, future developments of the study could consist of a clearer examination of affiliation-based closure role in community management contexts. It is to deeper understand which reasons determines the collaboration between organizations belonging to the same social foci, in terms of collective effectiveness and network culture: the significance nodes give to their collaboration activities, to their strategies, and to resources sharing (i.e. why and how much they are willing to collaborate, how interactions among smaller and bigger organizations work, which motivation can improve the creation of a collective identity, etc.).

By this, this kind of knowledge may be helpful for increasing social capital and intra/inter-organizational networks to a more strategic resources management, and to a wider adoption of pro-social behaviors. And for better realizing "bottom-up" policies directed to community empowerment, to community work, and to solve social problems.

References

Amati V., Lomi A., Mascia D., Pallotti F. (2019), "The Co-evolution of Organizational and Network Structure: The Role of Multilevel Mixing and Closure Mechanisms", *Organizational Research Methods*, July 30, pp. 1-34.

Anklam P. (2007), *Net Work. A Practical Guide to Creating and Sustaining Networks at Work and in the World*, Butterworth-Heinemann, Oxford.

Ansell C., Gash A. (2007), "Collaborative governance in theory and practice", *Journal of Public Administration Research and Theory*, 18, pp. 543-571.

Bellotti E. (2015), *Qualitative Networks. Mixed methods in sociological research*, Routledge, London-New York.

Bertin G., Fazzi L. (2010), *La governance delle politiche sociali in Italia*, Carocci, Roma.

Blumer H. (1969), *Symbolic Interactionism. Perspective and Methods*, Prentice Hall, Englewood Cliffs (NJ).

Bodin Ö., Nohrstedt D. (2016), "Formation and performance of collaborative disaster management networks: Evidence from a Swedish wildfire response", *Global Environmental Change*, 41, pp. 183-194.

Brailly J., Favre G., Chatellet J., Lazega E. (2016), "Embeddedness as a multilevel problem: A case study in economic sociology", *Social Networks*, 44, pp. 319-333.

Brennecke J., Rank O.N. (2016), "The interplay between formal project memberships and informal advice seeking in knowledge-intensive firms: A multilevel network approach", *Social Network*, 44, pp. 307-318.

Charmaz K. (2008), "*Grounded Theory as an Emergent Method"*, in S.N. Hesse-Biber, P. Leavy (*eds.*), *The Handbook of Emergent Methods*, Guildford, New York, pp. 155-170.

Christopoulos D. (2008), "The Governance of Networks: Heuristic or Formal Analysis? A Replay to Rachel Parker", *Political Studies*, 56 (2), pp. 475-481.

Contractor N., Wasserman S., Faust K. (2006), "Testing multi-theoretical multilevel hypothesis about organizational networks: An analytic framework and empirical example", *Academy of Management Review*, 31 (3), pp. 681-703.

Feld S.L. (1981), "The focused organization of social ties", *The American Journal of Sociology*, 86, pp. 1015-1035.

Frank O., Strauss D. (1986), "Markov Graphs", *Journal of the American Statistical Association*, 81, 395, pp. 832-842.

Granovetter M. (1992), "Problems of Explanation in Economic Sociology", in N. Nohria, R.G. Eccles (*eds.*), *Networks and Organizations: Structure, Form and Action*, Harvard Business School Press, Boston, pp. 25-56.

Hollway J., Koskinen J. (2016), "Multilevel embeddedness: The case of the global fisheries governance complex", *Social Network*, 44, pp. 281-294.

Jones C., Hesterley W.S., Borgatti S. (1997), "A General Theory of Network Governance: Exchange Conditions and Social Mechanisms", *Academy Management Review*, 22 (4), pp. 911-945.

Kazepov J. (2009), *La dimensione territoriale delle politiche sociali in Italia*, Carocci, Roma.

Kickert W.J.M., Klijn E.H., Koppenjan J.F.M. (1997), *Managing Complex Networks: Strategies for the Public Sector*, SAGE, London.

Krebs V., Holley J. (2006), *Building smart communities through network weaving*, www.orgnet.com.

Lazega E., Pattison P. (1999), "Multiplexity, Generalized Exchange, and Cooperation", *Organizations. Social Network*, 21, pp. 67-90.

Lomi A., Pallotti F. (2013), "How To Close a Hole: Exploring Alternative Closure Mechanisms in Interorganizational Networks", in D. Lusher, J.H. Koskinen, G.L. Robins (*eds.*), *Exponential Random Graph Models for Social Networks: Theory, Methods and Applications*, Cambridge University Press, Cambridge (UK), pp. 202-213.

Lomi A., Stadtfeld C. (2014), "Social Networks and Social Settings: Developing a Coevolutionary Vie", *Springer Fachmedien Wiesbaden*, 66, pp. 395-415.

Lusher D., Koskinen J.H., Robins G.L. (*eds.*) (2013), *Exponential Random Graph Models for Social Networks: Theory, Methods and Applications*, Cambridge University Press, Cambridge (UK).

McGlashan J., de la Haye K., Wang P., Allender S. (2019), *Collaboration in Complex Systems: Multilevel Network Analysis for Community-Based Obesity Prevention Interventions*, Scientific Reports.

McPherson M., Smith-Lovin L., Cook J. (2001), "Birds of a feather: Homophily in social networks", *Annual Review of Sociology*, 27, pp. 415-444.

Mische A., Pattison P. (2000), "Composing a civic arena: Publics, projects and social settings", *Poetics*, 27, pp. 163-194.

Molm L., Collett J., Schaefer D. (2007), "Building solidarity through generalized exchange: A theory of reciprocity", *American Journal of Sociology*, 113, pp. 205-242.

Mumford M.D., Antes A.L., Caughron J.J., Friedrich T.L. (2008), "Charismatic, ideological, and pragmatic leadership: Multi-level influences on emergence and performance", *The Leadership Quarterly*, 19 (2), pp. 144-160.

Provan K.G., Kenis P. (2008), "Modes of network governance: structure, management, and effectiveness", *Journal of Public Administration Research and Theory*, 18 (2), pp. 229-252.

Provan K.G., Milward H.B. (1995), "A Preliminary Theory of Interorganizational Network Effectiveness – A Comparative Study of 4 Community Mental Health Systems", *Administrative Science Quarterly*, 40 (1), pp. 1-33.

Provan K.G., Veazie M.E., Staten L.K., Teufel-Shone N.I. (2005), "The Use of Network Analysis to Strenghten Community Partnership", *Public Administration Review*, 65 (5), pp. 603-612.

Psaroudakis I. (2011), "Network Perspectives for Community Building", in A. Salvini, J.A. Andersen (*eds.*), *Interactions, Health and Community*, Pisa University Press, Pisa, pp. 205-215.

Raab J., Kenis P. (2009), "Heading toward a society of networks: empirical development and theoretical challenges", *Journal of Management Inquiry*, 18 (3), pp. 198-210.
Robins G., Bates L., Pattison P. (2011), "Network Governance and Environmental Management: Conflict and Cooperation", *Public Administration*, 89 (4), pp. 1293-1313.
Robins G., Pattison P., Kalish Y., Lusher D. (2007), "An Introduction to Exponential Random Graph (p*) models for Social Networks", *Social Networks*, 29 (2), pp. 173-191.
Salamon L.M., Anheier H.K. (1998), "Social Origins of Civil Society: Explaining the Nonprofit Sector Cross-Nationally", *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 9, pp. 213–248.
Salvini A. (2011), "*Dentro le reti. Forme e processi della Network Governance*", *Rivista Trimestrale di Scienza dell'Amministrazione*, 4, pp, 39-58.
Salvini A., Psaroudakis I. (2016), "Dynamics and Forms of a Network Solidarity. Towards an Innovative Practice of Networking between Volunteer Organizations", in A. Borghini, E. Campo (*eds.*), *Exploring the crisis: theoretical perspectives and empirical investigation*, Pisa University Press, Pisa, pp. 89-104.
Salvini A., Riccardo A., Vasca F., Psaroudakis I. (2020), "Inter-Organizational Networks and Third Sector: Emerging Features from two Case-Studies in Southern Italy", in G. Ragozini, M.P. Vitale (*eds.*), *Challenges in Social Network Research. Methods and Applications*, Springer, Cham, pp. 209-233.
Snijders T., Pattison P.E., Robins G.L., Handcock M.S. (2006), "New specifications for exponential random graph models", *Sociological Methodology*, 36, pp. 99-153.
Snijders T.A., Bosker R. (2012), *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modelling*, SAGE, London, 2nd edition.
Uzzi B. (1996), "The sources and consequences of embeddedness for the economic performance of organizations: the network effect", *American Sociological Review*, 61, pp. 674-698.
Vasca F., Riccardo A., Capuano G. (*eds.*) (2016), *Reti di periferia. Sistemi sociali virtuosi fra Terra di Lavoro e Terra dei Fuochi*, Aracne, Roma.
Vasca F., Tangredi D., Riccardo A., Iervolino R. (2016), "Modelli di costruzione di comunità tra coopetizione e cooperosità", in A. Salvini, *Dinamiche di comunità e servizio sociale*, Pisa University Press, Pisa, pp. 117-139.
Wagner A. (2000), "Reframing 'Social Origins Theory': The Structural Transformation of the Public Sphere", *Non Profit and Voluntary Sector Quarterly*, 29, 4, pp. 541-553.
Wang P., Robins G., Pattison P., Lazega E. (2013), "Exponential random graph models for multilevel networks", *Social Networks*, 35 (1), pp. 96-115.
Wang P., Sharpe K., Robins G., Pattison P. (2009), "Exponential random graph (p*) models for affiliation networks", *Social Networks*, Elsevier, 31 (1), pp. 12-25.
Wang R., Tanjasiri S.P., Palmer P., Valente T.W. (2016), "Network Structure, Multiplexity, and Evolution as Influences on Community-based Participatory Interventions", *Journal of Community Psychology*, 44 (6), pp. 781-798.

Wang W. (2015), “Exploring the Determinants of Network Effectiveness: The Case of Neighborhood Governance Networks in Beijing”, *Journal of Public Administration Research and Theory*, 26 (2), pp. 375-388.
Wasserman S., Pattison P.E. (1996), “Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p*”, *Psychometrika*, 61, pp. 401-425.
Wellman B., Berkovitz (1988), *Social Structure. A Network Approach*, Cambridge University Press, Cambridge (UK).
Zappa P., Lomi A. (2015), “The Analysis of Multilevel Networks in Organizations: Models and Empirical Tests”, *Organizational Research Methods*, 18 (3), pp. 542-569.
Zappa P., Lomi A. (2016), “Knowledge Sharing in Organizations: A Multilevel Network Analysis”, in E. Lazega, T.A.B. Snijders (*eds.*), *Multilevel Network Analysis for the Social Sciences. Theory, Methods and Applications*, Method Series 12, Springer, Cham, pp. 333-353.

Part II
Empirical and Substantive Findings

 ISBN 9788835124603

5. *Social network analysis for welcome policies: an empirical study from small Italian municipalities*

by Ilaria Marotta[*], Carmela Bernardo[**], Angelo Moretti[***], Francesco Giangregorio[****], Francesco Vasca[*****]

1. Introduction

Interorganizational networks are characterized by interesting emerging behaviors which can be studied through the theoretical construct of the Social Network Analysis (SNA) (Ebers, 2015; Barabasi, 2016; Wasserman and Faust, 1994). In particular, SNA has been proposed in the literature for the analysis of interactions among organizations that share information and resources for counteracting certain social problems, or for promoting innovations in social, educational and environmental interventions (Raab and Kenis, 2009; Provan *et al.*, 2005; Salvini, 2011).

The strategy of analysis proposed in this work adopts the SNA approach for studying interorganizational networks aimed at the development of cooperatives dedicated to local welcome policies. The proposed network analysis derives from an empirical study carried out within the project entitled I Piccoli Comuni del Welcome (The small municipalities of welcome) activated in the provinces of Benevento and Avellino in Southern Italy and supported by the Italian foundation Fondazione CON IL SUD[1]. The main objective of

[*] Department of Social Sciences, University of Naples Federico II – former University of Sannio, ilaria.marotta@unina.it.

[**] Department of Engineering, University of Sannio, cbernardo@unisannio.it.

[***] Consorzio Sale della Terra, Benevento, amorettibn@gmail.com.

[****] Consorzio Sale della Terra, Benevento, direzioneamministrativasdt@gmail.com.

[*****] Department of Engineering, University of Sannio, vasca@unisannio.it.

[1] Fondazione CON IL SUD was launched in November 2006 as a result of the alliance between bank foundations and Italian third sector organizations to foster social infrastructure in Southern Italy, focusing on building up and qualifying the intangible structures to promote development. The project I Piccoli Comuni del Welcome was approved as part of an open

the project is to launch the constitution of so called Community Cooperatives (CoCo) – proposed by citizens – whose goal is to improve the quality of life in small municipalities where national Systems of PRotection for Asylum seekers and Refugees (SPRAR) have been previously activated. Building a CoCo is expected to contribute to the innovation of the social and economic dynamics of the territory thus promoting local development.

This study intends to analyze the networking actions emerging among the CoCo and other territorial entities involved in welcome policies and collaborating for the cooperative development. The whole-network strategy identifies the boundaries of the CoCo network by selecting all actors involved in the cooperative construction goal. Two attributes for each edge of the network are considered: mutual knowledge and trust. The values of the attributes are obtained by means of dedicated questionnaires submitted to the network nodes. The analysis of centralities and clusters provides useful indicators for monitoring and evaluating the governance of the network of interacting organizations and for identifying possible directions for improving their relationships.

2. Theoretical framework

In Italy, the integration policies for migrants have been traditionally characterized by strong territorial differentiations (Zincone, 2006). Parallel to migration policies, in the same country there exists the SPRAR reception system which focuses on the network of authorities for the development of territorial projects for the inclusion of asylum seekers and refugees in local communities. This type of interventions requires substantial investments and often provokes controversies among the population (Ambrosini, 2000). For this reason, the welcome perspective is a privileged observation point for highlighting the strategies of action of local actors in the face of dilemmas that are not easy to solve. Searching for possible solutions of the issues indicated above has reinforced in recent years the involvement of volunteering and third sector organizations in welcome projects.

Furthermore, in a scenario of redefinition of the relationship between "center" and "periphery" (Della Porta, 1999) and calibration of policies towards the territorial dimension, local authorities have started to seek integration through a system of partnerships with different subjects – including the local administration – to plan this type of policy. In this framework, the

call for proposals concerning immigrants, their socio-economic integration and related social emergencies proposed by third sector organizations with a maximum total of 2.5 million euros.

subsidiarity principle attributes to third sector organizations the role of co-agents in the definition and implementation of policies (Bifulco, 2005; Bifulco, 2008; Donolo, 2005; Paci, 2008) and provides the creation of a network between different actors and territories whose main objective is the management of the common good (Corbisiero, 2009).

The importance of the diffusion of territorial subsidiarity (Endo, 1994; Moreno, 2007) highlights how in specific territories the subjects of the third sector are mobilized and integrated according to heterogeneous logics, methods and practices. This action of bringing together institutional and third sector actors allows the construction of networks with public connections through a social and political coordination (Piselli, 2010). The result of these policies is no longer linked to the performance of each actor but to the synergistic action of several interdependent subjects included in the network.

In light of this, the image of the network has in recent years taken on a strong analytical value in describing the form that underlies the governance structures of social policies, since, in Italy, the Law for the reform of social services n. 328/00 promoted the transition to decentralized governance models (Fedele, 2005). This means that the system of social policies is based on a collaboration that brings together different levels of government, both public and private subjects (Pavolini, 2003). It is evident, based on what has been said so far, that the concept of interorganizational network represents a fundamental point of view for studying this phenomenon.

"Social network analysis views interorganizational networks as a set of linkages (e.g., resource, friendship, informational ties) among a set of actors (individuals, groups, or organizations)" (Ebers, 2015), in order to achieve a common result (output) or to jointly produce an expected emergent behaviour (outcome). The interorganizational network refers to a metaphor of the interdependent nature of organizations because interactions occur beyond the boundaries of the organization itself. An interesting classification of interorganizational networks considers two different classes (Raab & Kenis, 2009): the first type refers to a structure of relationships of informal interdependence between collective social actors that "emerges" from their dyadic interactions, without further specification of objectives, timing and ways of "being together"; the second type is more concerned with network governance, whereby the systems of interdependence between collective entities are intentionally created to achieve certain goals. The study of this work refers to the second type of networks because the organizations involved collaborate with the CoCo for local welcome policies.

A crucial question when dealing with interorganizational networks is how the effectiveness of the network governance can be assessed (Wang, 2016;

Robins, Bates & Pattison, 2011; Christopulos, 2008). This effectiveness can be defined as "the attainment of positive network-level outcomes that could not be achieved by individual organizational participants acting independently" (Provan & Kenis, 2008). The assessment of network effectiveness must necessarily take into consideration a plurality of aspects, such as the impact of network activities in the served community, the role of individual members and the nature of the interactions between them, the evaluation of the stakeholders, and, obviously, the characteristics of the structure of the interaction network (Provan & Milward, 1995).

The SNA methodological perspective has been shown to be appropriate and comprehensive for studying the effects of the network structure on governance and outcomes, as well as in supporting the management of these governance processes (Cross, Borgatti & Parker, 2002; Salvini, 2011; Salvini *et al.*, 2020). SNA makes available a set of methodological tools and techniques to verify if and how the structural configuration of those relationships generates effects both on the governance of the network and on its outcomes (Provan *et al.*, 2005; Provan, Fish & Sydow, 2007). This could encourage the involvement and empowerment of collective actors operating at the level of local communities (Wang *et al.*, 2016; Neal, 2014): the intuition on the importance of working together and sharing resources can be anchored through SNA into a solid and validated conceptual and theoretical framework.

3. Network definition and analysis

The analysis of the interorganizational network of the local actors involved in the CoCo development is based on the SNA theoretical framework. In this section, the procedures for the identification of the territorial actors (the nodes of the digraph), the definition of the mutual knowledge and trust relationships among the organizations (the attributes of each edge) and the evaluation of the intensity of these relationships (the weights of the edges) are described.

The nodes of the digraph are selected by considering the organizations (not individual people) involved in the CoCo development, i.e. the goal of the network. A preliminary list of actors has been suggested by a privileged witness. Specific interviews to these nodes allowed the identification of more actors of interest for the analysis, by using a snowball sampling approach, so that the network was representative not only of the actors strictly related to the CoCo but of the whole territorial relational system related to the welcome policies.

In order to understand in depth, the structure and functioning of these welcome networks in the specific territorial contest, two types of relationships between pairs of nodes have been considered: mutual knowledge and trust. The resulting multiplex network can be represented as a digraph where the weights of each edge are the measurements of the two types of relationships between the connected nodes (Dickison, Rossi & Magnani, 2016).

The mutual knowledge reflects the current configuration of the familiarity between the actors; on the other hand, the trust, intended as a propensity for collaboration, helps to understand the role that the different actors can assume within the network and possible obstacles for future collaborations.

A questionnaire was submitted to each actor identified as participating in the welcome policies in the specific territory considered. Some preliminary information on the characteristics of the organization and comments on the problematic aspects for the creation of a network in the territory were firstly recorded through the interview. In order to identify the relationships of mutual knowledge, the question of the questionnaire was: "How much do you know the other organization?", by listing the other nodes of the network. Each answer could be chosen as one of the following four levels of intensity:

- 0: I do not know the entity;
- 1: I know little about the entity (for example, I know just the name);
- 2: I know the entity (for example, I know its seat, I had sporadic contact with some of its members, I follow its social networks);
- 3: I know the entity very well (for example, I interact very frequently with it).

The question regarding the relationship of trust was: "How much would you collaborate with the entity?" and the possible options for the answer were:

- 0: I do not trust the entity, or I do not know;
- 1: I have little inclination to collaborate with the entity;
- 2: I have a sufficient propensity for collaborating with the entity;
- 3: I have a good propensity for collaborating with the entity.

The mutual knowledge and trust attributes were combined with two different selections of the nodes thus generating the local network and the territorial network. The local network is made by all organizations which operate in (and only in) the specific area of the CoCo. The territorial network includes all entities that collaborate with the CoCo, regardless of their geographic location and area where the activities are carried out. The comparison between the territorial and the local networks allows one to identify the most strategic entities for inducing the cooperation among the actors as well as the local entities on which one should invest to increase the resilience potentialities of the network itself.

The network has been analyzed under the perspective of local welcome policies. The global digraph indices (density, reciprocity and diameter) provide information with respect to the diffusion of the local interactions. The indices calculated on the individual nodes (in-degree, betweenness, hub and authority centralities) allow one to identify the network hubs (Kleinberg, 1999).

4. Data collection

The case study considered is the CoCo called Tralci di Vite which was established in 2017 in the small municipality of Chianche under the activities of the project I Piccoli Comuni del Welcome.

The principal investigator of the project is the social cooperative Il Melograno which submitted the project by responding to a call published by the Italian foundation Fondazione CON IL SUD in 2017. Il Melograno is part of a network of cooperatives and associations that share a systemic action for the development of the territory with the diocesan Caritas of Benevento starting from the promotion of well-being and youth and female entrepreneurship, with particular attention to social disadvantage. The main goal of the project is to launch the constitution of 10 community cooperatives in some municipalities in the provinces of Benevento and Avellino in Sothern Italy, in particular where SPRARs were already activated. These two provinces are experiencing a profound social crisis due to negative migratory balances with an extraordinary emptying of the territories and an increasingly pervasive abandonment of the usable agricultural area (Bock, 2016; Svimez, 2019).

The project pursues the general objective of establishing low-profit and non-profit business practices with high repercussions in terms of improving the quality of life for beneficiary immigrants and unemployed natives and with an intention to return to the native communities, through the establishment of these mixed cooperatives and their network coordinated by the social cooperative Il Melograno. The main activities of the cooperatives are social agriculture, wine and oil market, tourism, handicraft, installation and maintenance of photovoltaic systems. These activities are expected to lead to stable job positions for 160 people, including 100 immigrants, in 5 years.

The characteristics of the project make it replicable in other Italian rural areas at risk of depopulation that intend to promote welcome experiences based on the SPRAR model. Considering the number of municipalities at risk of abandonment, the repeatable potentialities of the project are widespread.

In Italy, the small municipalities – under 5 thousand inhabitants – are 73% of the total, and insist on 54% of the territory, with a population that stands at 26% on the national average (De Blasio, Giorgione & Moretti, 2018).

The CoCo Tralci di Vite was the first cooperative created under the project I Piccoli Comuni del Welcome. It was born on 25 September 2017 in the municipality of Chianche with 7 founders, including operators and beneficiaries of the SPRAR. The cooperative is currently based on agricultural activities, social and health services, recovery of abandoned land, training courses in agriculture, maintenance of public and private green areas and management of a local market.

The network of the CoCo is composed of 42 organizations, including the cooperative itself, which have been classified in 10 different categories. The distribution among the categories and the answers to the questionnaire received are shown in Tab. 1.

Tab. 1 – Entities which answered to questionnaire, by types of organization

Acronym	*Category*	*Total*	*Answers*
AS	Association	3	3
CN	Consortium	1	1
CO	Committee	3	3
CP	Cooperative	11	11
RB	Religious body	4	3
TB	Training body	2	2
CM	Company	7	1
IS	Institution	5	1
SU	School/University	3	3
OT	Others	3	0
	Total	42	28

The questionnaires survey was done from July 2018 till March 2019, by obtaining 67% of the total expected responses from territorial entities and 46% of responses from local entities.

Among the 28 actors who participated in the survey, most of them represent organizations already rooted in the territory. In fact, 10 organizations have been operating for more than 20 years in the area, while 5 organizations have been present from 11 to 20 years. The other entities have a more recent presence in the territory: 4 subjects work in the area from 6 to 10 years, 6 from 2 to 5 years. Finally, 3 are recent organizations, born in the last reference year.

Most respondents deal with the protection of civil rights, citizenship, immigrants, minors and women, to follow socio-health services such as care for the elderly, disabled, destitute and cultural services such as protection of cultural heritage, training and research.

As regards the number of resources (volunteers, operators, employees) committed in the last year of the organization activities, 11 of the respondent subjects indicated more than 20 resources. Following this, 6 organizations indicated between 2 and 5 resources; while only 2 subjects indicated 0 or just 1 resource.

Finally, each respondent was asked to indicate the level of problematicity for some critical aspects in creating a network within the reference territory. Tab. 2 shows the corresponding responses. The most problematic aspects of creating a network within the area of competence are, in order of importance: the lack of awareness of the importance of the network; the difficulty in putting together tangible and intangible resources; the fragmentation of interorganizational activities and self-reference/individuality.

Tab. 2 – Number of responses from the organizations for each problematic aspect in the creation of a network in the area of competence (rows) and for each level (columns)

	Not at all problematic	*Unproblematic*	*Neutral*	*Quite problematic*	*Highly problematic*
Coordination of activities	5	6	7	7	3
Diversity of aims/ objectives	5	10	4	4	5
Lack of confidence	7	10	6	3	2
Difficulties in combining tangible and intangible resources	7	5	6	9	1
Self-referentiality and individuality	5	6	6	8	3
Fragmentation of inter-organizational relations	3	6	5	10	4
Inability to exchange information	4	8	5	8	3
Lack of awareness of the importance of the network	3	3	7	13	2
Inability to assume responsibility	4	5	9	8	2

5. Main findings for the territorial and local networks

In this section, the territorial network consisting of all identified actors and its sub-network including only the organizations established in Chianche, which is called local network, are analyzed. The weights of the edges outgoing from the nodes corresponding to entities which did not answer the questionnaire are set to zero. Moreover, the analysis reserved only for the entities that responded to the questionnaire is also carried out. The results show the centrality of the cooperative Tralci di Vite which can be considered as a bridging actor between local and territorial entities.

5.1. The territorial network

The territorial network consists of all actors involved in the CoCo constitution, directly and indirectly. The answers of all entities have determined the adjacency matrices of the two weighted digraphs obtained by considering all 42 nodes and two different attributes (mutual knowledge and trust) for the edges. The corresponding network diagrams are shown in Fig. 1 and Fig. 2. The gray intensity of each node is proportional to the in-degree centrality, i.e. the number of connections entering that node by taking into account the weights of the edges.

The subnetwork obtained by considering the only 28 nodes corresponding to the actors who responded to the questionnaire has been extracted. The global measures of this subnetwork together with those of the whole territorial network are reported in Tab. 3.

The number of edges and, consequently, the density in the case of the trust attribute is greater than the case of the mutual knowledge attribute. Then, one could say that there is a propensity to collaboration even if some entities do not know each other. Obviously, the density has higher values for the network consisting of 28 nodes, since in its construction the entities which did not provide answers, i.e. those nodes without outgoing edges in the network consisting of all 42 nodes, are excluded.

As regards the reciprocity (percentage of reciprocal connections), its value is similar for both mutual knowledge and trust attributes. The values assumed by reciprocity are quite high, so perceptions are sufficiently reciprocated among the nodes.

Fig. 1 – Diagram of the territorial network obtained by considering the mutual knowledge attribute. The darker the gray tone of the node, the greater its in-degree centrality

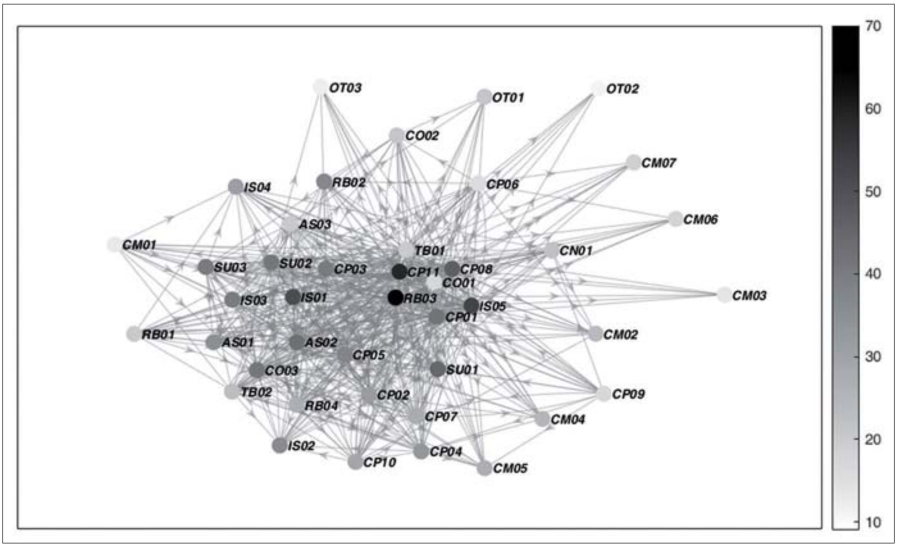
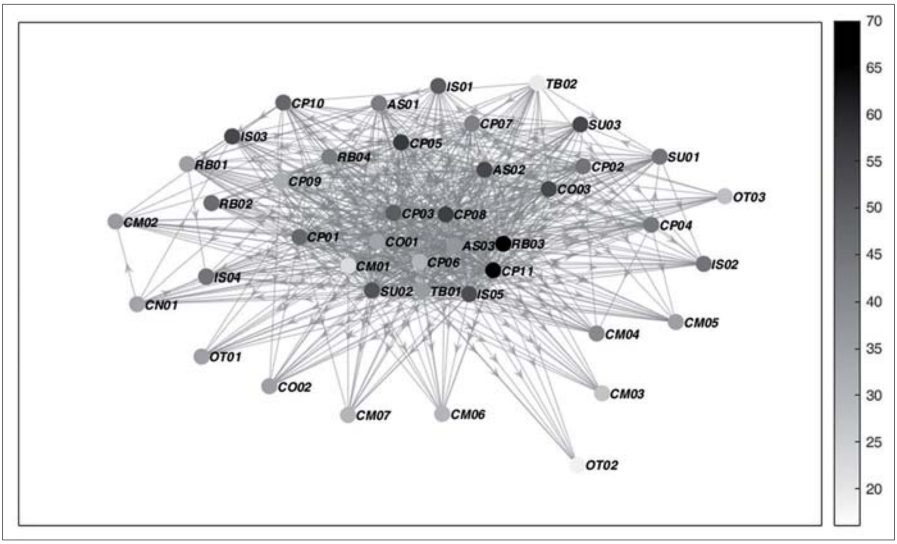


Fig. 2 – Diagram of the territorial network obtained by considering the trust attribute. The darker the gray tone of the node, the greater its in-degree centrality



Tab. 3 – Global measures of the territorial networks with 42 nodes (second and third columns) and the only 28 nodes who responded to the questionnaire (fourth and fifth columns)

	Network with 42 nodes		*Network with 28 nodes*	
Measures	*Mutual knowledge*	*Trust*	*Mutual knowledge*	*Trust*
Edges	604	689	423	480
Density	0.35	0.40	0.56	0.63
Reciprocity	0.46	0.45	0.71	0.69
Diameter	1.33	0.83	1.33	0.83

In order to analyse the relationships among the nodes in terms of their distances, two more edges attributes have been considered: knowledge distance and trust distance. The edges weights are obtained as the reciprocal of the level indicated in the corresponding answer. For example, if the node A knows the node B with a level equal to 2, the edge attribute corresponding to the knowledge distance from A to B has a weight equal to 0.5. In particular, for the null entries of the original adjacency matrices, no edges have been considered in the corresponding distance attribute too. The resulting diameter for the trust distance (0.83 in Tab. 3) is smaller than that for the mutual knowledge distance (1.33 in Tab. 3), which is coherent with the density results.

The relevant and strategic nodes in the networks can be identified by evaluating the distribution of the network centralities among the nodes.

The in-degree distribution for the mutual knowledge attribute in the territorial network is shown in Fig. 3. The nodes RB03 (religious body) and CP11 (cooperative) are those with the highest in-degree centrality (70 and 59, respectively), i.e., they are the most known entities in the network. All the other nodes have an in-degree centrality within the interval [11, 53].

By computing the betweenness centrality, the nodes RB03 (betweenness equal to 230) and CP11 (betweenness equal to 193) are also the most strategic in the network with respect to the connectivity features, allowing the shortest path connections between actors who do not know much about each other (all the other nodes have a betweenness less than 100). The nodes RB03 and CP11, together with CP02, are also the most strategic in terms of hub and authority centrality for the territorial network.

Nodes clustering obtained by considering the in-degree and the betweenness centralities for the mutual knowledge attribute can be deduced from the dendrograms shown in Fig. 4. As regards the in-degree, when the distance level is equal to 11, the agents group in 3 clusters with 15, 25 and 2 members, respectively. Note that the cluster with 2 elements (those furthest to the right in the dendrogram) consists of the nodes RB03 and CP11 which have the highest in-degree centrality. At a distance level around 20, the two most numerous clusters merge in a single one. Finally, at a distance level of 36 all nodes of the network are grouped in a single cluster. The betweenness dendrogram on the right of Fig. 4 clearly shows that most nodes connect into a cluster at a very low distance level because of the low values of betweenness for almost all nodes of the network.

Fig. 3 – In-degree distribution in the territorial network by considering the mutual knowledge attribute

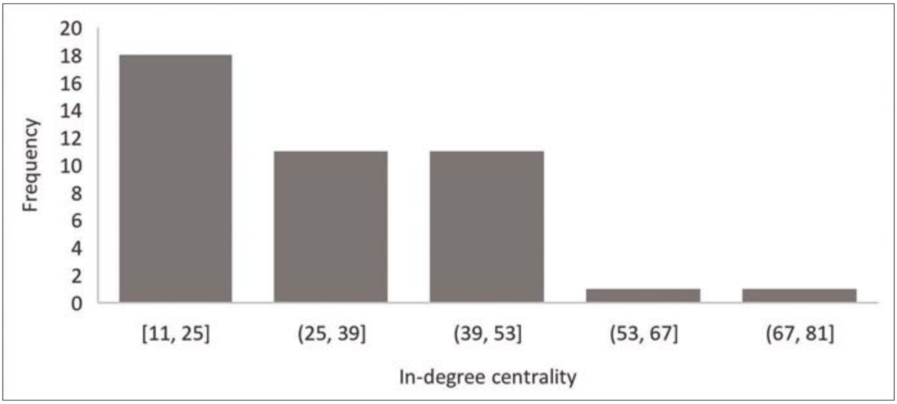
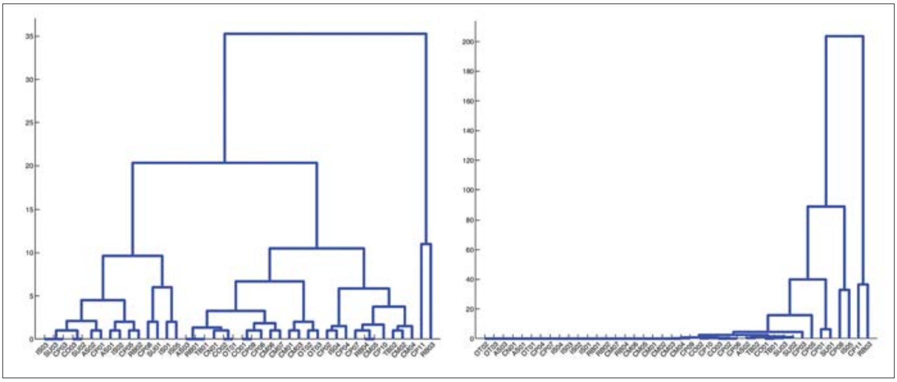


Fig. 4 – Dendrograms of the in-degree (left) and betweenness (right) measures for the territorial network with the mutual knowledge attribute



In synthesis, the network analysis shows that the two nodes which play a strategic role in terms of in-degree and betweenness centralities are the territorial entities which are the principal promoters of the CoCo.

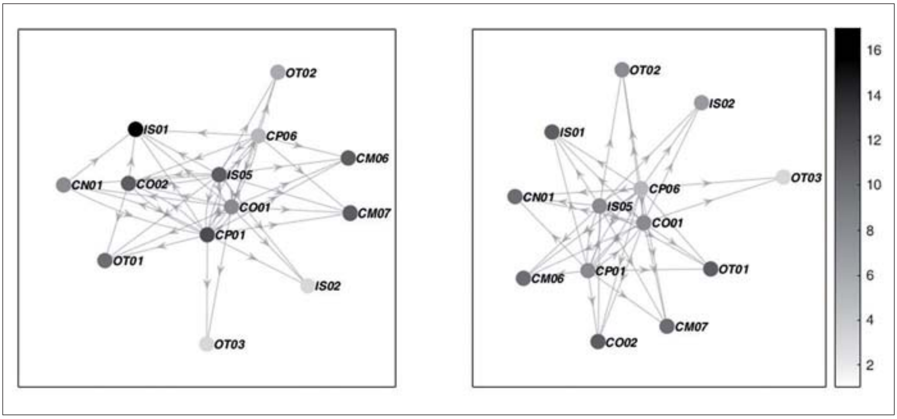
The same conclusions about the most strategic nodes of the network are obtained by considering the trust attribute and by restricting the analysis to the 28 entities which responded to the questionnaire.

5.2. The local network

The local network is obtained by selecting the 13 nodes belonging to the municipal territory of Chianche and the edges among the corresponding nodes. Note that this local network cannot be obtained as an ego-centered network extracted from the territorial network by choosing any of its nodes as the ego. The evaluation of the in-degree and betweenness centralities identifies an important local actor which represents a bridge between the local entities and the others.

The digraphs generate by considering the two edges attributes (mutual knowledge and trust) in the local network are shown in Fig. 5.

Fig. 5 – Diagram of the local network obtained by considering the mutual knowledge (left) and the trust (right) attributes. The darker the gray tone of the node, the greater its in-degree centrality



The local networks global measures reported in Tab. 4 show that the density for the mutual knowledge attribute is greater than that for the trust attribute. A possible interpretation is that a high level of mutual knowledge brings a reduction of trust.

The computation of the betweenness centrality in the local network shows that the node CP01 has a central role for the mutual knowledge. Moreover, the entities CO02 and IS01 represent the authorities, while CO01, CP01 and

IS05 are the hubs of the local network. These considerations are valid for both networks consisting of 13 and 6 nodes.

Tab. 4 – Global measures for the two local networks consisting of the 13 local nodes (second and third columns) and the only 6 nodes of them who responded to the questionnaire (fourth and fifth columns)

	Network with 13 nodes		*Network with 6 nodes*	
Measures	*Mutual knowledge*	*Trust*	*Mutual knowledge*	*Trust*
Edges	47	43	21	19
Density	0.30	0.28	0.70	0.63
Reciprocity	0.34	0.23	0.76	0.53
Diameter	2	1	1.5	0.67

In synthesis, the analysis of the networks has shown that the node CP01, which is the recently founded CoCo in the territory of Chianche, has an important connecting position for the nodes of the local network and has centralities measures sufficiently high in the territorial network. Then, the node CP01 can be considered as a bridging actor between local and non local entities.

6. Conclusions

In this work, SNA has been used for studying the collaborations among organizations involved in the construction of a community cooperative. The analysis is supported by an empirical study carried out in a small municipality in Southern Italy. The connections among the nodes, based on mutual knowledge and trust, have been evaluated through the submission of questionnaires which determined the adjacency matrices for the weighted digraphs representing the interorganizational networks. A local sub-network constructed by considering the actors which operate only on the territory of interest has been considered too. The analysis of the networks leads to the interpretation that in the small municipality trust seems to be reduced by the high mutual knowledge. The higher reciprocity indicates a marked propensity for collaboration, which is based on a sharing of common objectives. Moreover, the centralities measures for the node representing the local cooperative demonstrate the importance of that entity as a bridging actor between local and non local entities.

Future work will concern a dynamic assessment of the networks involved in the establishment of community cooperatives aimed at highlighting the variation in the characteristics of the network generated over time as a result of the collaborations among the organizations involved.

References

Ambrosini M. (2000), "Immigrati e lavoro in Italia. Le dimensioni culturali di un processo economico [Immigrants and work in Italy. The cultural dimensions of an economic process]", *Aggiornamenti sociali*, 2, pp. 145-157.

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

Bifulco L. (2008), *Gabbie di vetro* [*Glass cages*], Mondadori, Milano.

Bifulco L. (a cura di) (2005), *Le politiche sociali. Temi e prospettive emergenti* [*Social policies. Emerging themes and perspectives*], Carocci, Roma.

Bock B.B. (2016), "Rural marginalisation and the role of social innovation; a turn towards nexogenous development and rural reconnection", *Sociologia Ruralis*, 56 (4), pp. 552-573.

Christopulos D. (2008), "The Governance of Networks: Heuristic or Formal Analysis? A Replay to Rachel Parker", *Political Studies*, 56 (2), pp. 475-481.

Corbisiero F. (2009), *Lavoro flessibile e forme contrattuali nonstandard nel terzo settore [Flexible work and non-standard contractual forms in the third sector]*, FrancoAngeli, Milano.

Cross R., Borgatti S.P., Parker A. (2002), "Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration", *California Mangement Review*, 44 (2), pp. 25-46.

De Blasio N., Giorgione D., Moretti A. (2018), *L'Italia che non ti aspetti* [The Italy you do not expect], CittàNuova, Roma.

Della Porta D. (1999), *La politica locale. Potere, istituzioni, attori tra centro e periferia [Local politics. Power, institutions, actors between the center and the periphery]*, il Mulino, Bologna.

Dickison M.E., Rossi L., Magnani M. (2016), *Multilayer Social Networks*, Cambridge University Press, Cambridge.

Donolo C. (2005), "Dalle politiche pubbliche alle pratiche sociali nella produzione di beni pubblici? Osservazione su una nuova generazone di policies [From public policies to social practices in the production of public goods? Observation on a new generation of policies]", *Stato e mercato*, 25 (1), pp. 33-66.

Ebers M. (2015), "Interorganizational Relationships and Networks", *International Encyclopedia of the Social & Behavioral Sciences*, 2, pp. 621-625.

Endo K. (1994), "The Principle of Subsidiariety: From Johannes Althusius to Jacques Delors", *Hokkaido Law Review*, 44 (6), pp. 553-652.

Fedele M. (2005), "Le basi istituzionali del welfare locale. Processi di fissione e fusione prima e dopo la 328 [The institutional bases of local welfare. Processes of fission and fusion before and after the 328]", *La Rivista delle Politiche Sociali*, 2, pp. 89-102.

Kleinberg, J.M. (1999), "Authoritative sources in a hyperlinked environment", *Journal of the ACM* (*JACM*), 46 (5), pp. 604-632.

Moreno L. (2007), "Europeanization, Territorial Subsidiarity and Welfare Reform", *Regional & Federal Studies*, 17 (4), pp. 489-497.

Neal Z.P. (2014), "A Network Perspective on the Process of Empowered Organizations", *American Journal of Community Psychology*, 53 (3-4), pp. 407-418.
Paci M. (2008), *Welfare locale e democrazia partecipativa [Local welfare and participatory democracy]*, il Mulino, Bologna.
Pavolini E. (2003), *Le nuove politiche sociali. I sistemi di welfare tra istituzioni e società civile* [*The new social policies. The welfare systems between institutions and civil society]*, il Mulino, Bologna.
Piselli F. (2010), "Partecipazione sociale e processi decisionali nel welfare territoriale del Mezzogiorno [Social participation and decision-making processes in the territorial welfare of the South]", *La Rivista delle Politiche Sociali*, 3, pp. 257-282.
Provan K.G., Fish A., Sydow J. (2007), "Interorganizational Networks at the Network Level: A Review of the Empirical Literature on Whole Networks", *Journal of Management*, 33 (6), pp. 479-516.
Provan K.G., Kenis P. (2008), "Modes of Network Governance: Structure, Management, and Effectiveness", *Journal of Public Administration Research and Theory*, 18 (2), pp. 229-252.
Provan K.G., Milward H.B. (1995), "A Preliminary Theory of Interorganizational Network Effectiveness – A Comparative Study of 4 Community Mental Health Systems", *Administrative Science Quarterly*, 40 (1), pp. 1-33.
Provan K.G., Veazie M.A., Staten L.K., Teufel-Shone N.I. (2005), "The Use of Network Analysis to Strengthen Community Partnership", *Public Administration Review*, 65 (5), pp. 603-613.
Raab J., Kenis P. (2009), "Heading Toward a Society of Networks: Empirical Development and Theoretical Challenges", *Journal of Management Inquiry*, 18 (3), pp. 198-210.
Robins G., Bates L., Pattison P. (2011), "Network Governance and Environmental Management: Conflict and Cooperation", *Public Administration*, 89 (4), pp. 1293-1313.
Salvini A. (2011), "Dentro le reti. Forme e processi della Network Governance [Inside the networks. Forms and processes of Network Governance]", *Rivista Trimestrale di Scienza dell'Amministrazione*, 4, pp. 39-58.
Salvini A., Riccardo A., Vasca F., Psaroudakis I. (2020), "Inter-Organizational Networks and Third Sector: Emerging Features from Two Case Studies in Southern Italy", in G. Ragozini, M. Vitale (*eds.*), *Challenges in Social Network Research. Lecture Notes in Social Networks*, Springer, New York, pp. 209-233.
Svimez (2019), *Rapporto sull'economia del Mezzogiorno* [*Report on the economy of the South*], retrieved from: http://lnx.svimez.info/svimez/il-rapporto/.
Wang W. (2016), "Exploring the Determinants of Network Effectiveness: The Case of Neighborhood Governance Networks in Beijing", *Journal of Public Administration Research and Theory*, 26 (2), pp. 375-388.
Wang R., Tanjasiri S.P., Palmer P., Valente T.W. (2016), "Network Structure, Multiplexity, and Evolution as Influences on Community-based Participatory Interventions", *Journal of Community Psychology*, 44 (6), pp. 781-798.

Wasserman S., Faust K. (1994), *Social Network Analysis – Methods and Applications*, Cambridge University Press, Cambridge.
Zincone G. (2006), "The making of immigration and immigrant policies", *Journal of Ethnic and Migration Studies*, 32 (3), pp. 347-375.

6. *"Suite of roles" as a driver of core-periphery patterns in water resource governance networks*

by Karen I. Trebitz*, Manoj K. Shrestha**

1. Introduction

Water basins are complex social-ecological systems, spanning political, legal, jurisdictional, socio-economic, geographic and biophysical boundaries (Bodin *et al.*, 2011). No single player has sole authority in water governance. Rather, it involves a wide variety of organizational players with differing agendas, resources, expertise, and authority. Further, water resource governance requires an adaptive process that addresses both place-based needs of water users and ecological challenges in such complex systems (Folke *et al.*, 2005). Network-weaving (Krebs & Holley, 2006), or reaching out to others to build and maintain networks of relationships and collaborations, is considered fundamental to solving the "perpetual crisis," of water policy (Scholz & Stiftel, 2005, p. ix).

Social networks are patterns of relationships among a finite set or sets of actors, which can be presented graphically and are evaluated empirically. Network structures, or patterns of connections, emerge as actors enter into or maintain relationships with others. Collective actions in resource governance result in polycentric arrangements, consisting of independent agencies and organizations in all sectors and government that work to achieve common goals (Ostrom, 2010).

One of the less studied network structures is the core-periphery network, with tight centers of densely connected actors (Yang & Leskovec, 2014). Actors in the peripheries are relatively close only to the well-connected central actors, but are only sparsely connected to other peripheral nodes (Borgatti &

* Department of Water Resources, University of Idaho, treb6275@vandals.uidaho.edu.
** Department of Politics and Philosophy, University of Idaho, mks@uidaho.edu.

Everett, 2000; Csermely *et al.*, 2013). Core periphery patterns are not uncommon, and have been observed in collaborative water resource management situations, such as in the St. Joe/St. Maries River sub-basin of Idaho's Lake Coeur d'Alene (the pilot for the current study), and the watershed governance sub-networks in Vermont's Lake Champlain Basin (Scheinert *et al.*, 2015).

Moreover, little research has been done to examine the drivers of core-periphery structures observed in resource governance networks (Smedstad & Gosnell, 2013; Kapucu *et al.*, 2014; Scheinert *et al.*, 2015; Jasny & Lubell, 2015). Comparative studies of social networks across multiple settings have been rare – due, in part, to a paucity of data sets (Steven Scheinert; Lorien Jasny, personal communications). In this study, we focus on drivers of the core-periphery structures observed in 18 water governance networks related to fisheries and/or water quality in five large reservoir basins of the greater Columbia River Basin. Focus questions in this paper are: *What actor attributes lead to core-periphery structures across the five lake basins? By implication, what attributes lead actors to be central in the network?*

Resource dependence theory posits that actions of organizations revolve around important resources (Pfeffer & Salancik, 1978/2003). Organizational actors in water governance seek services or capacities that further their goals in the basin. We propose that centrality of an actor in a network is not based on *who an actor is* in terms of a formal entity or its mission, but is a function of how many capacities or roles an actor fulfills in a given network – *what an actor does*. In this study, we developed a panel of actor capacities as a *suite of roles* composed of: *resources*, comprised of data, expertise, funding, and political support; *formal roles*, consisting of dam & resource tenure, regulatory jurisdiction, legal precedent; and *informal roles*, comprised of collaborations, communications, and policy entrepreneurship.

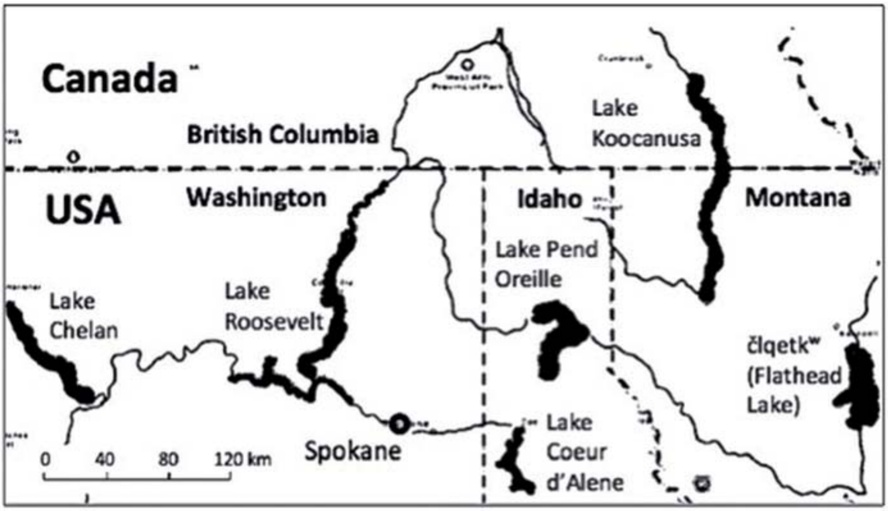
We used mixed methods research for this study that included collection and analyses of quantitative and qualitative data (Onwuegbuzie & Collins, 2007). Evaluating results from multiple techniques can result in more complete findings than the individual method can provide (Haunss *et al.*, 2017). We evaluated our proposition by combining methods in conventional statistics and social network analysis in a stepwise process. We began with scoping to develop a comprehensive list of possible actors involved in water resource governance in each basin. This process involved using archival data (Wasserman & Faust, 2009; Lubell & Lippert, 2011; Scheinert *et al.*, 2015), organizations' websites, and information supplied via emails and personal communications with basin actors. We then developed and distributed surveys to organizations identified in the list to be filled by representative officials responsible for the management of fisheries and water quality for each basin. Survey data were used to populate network matrices. Items constituting three role types – resource roles, formal roles and informal roles – for actors, were identified by using survey responses and planning documents, and by talking to water management practitioners in the study area. After visualizing the network graphs, we used Ucinet (Borgatti, Everett & Freeman, 2002) to test the presence of core-periphery network structures and the "statnet" package in R to perform quadratic assignment procedure (QAP) correlations to compare similarity of network patterns across basins based on the actor types. Finally, Pearson's correlations and Poisson regressions were used to evaluate the relationship between the suite of roles and centrality of actors. The results obtained in each step were consistent and supported our proposition as expected.

2. The study – setting and data collection

2.1. The study setting

In this study we focus on five large reservoir basins of the greater Columbia River Basin (Fig. 1).

Fig. 1 – Locations of the reservoir basins in this study, from west to east: Lakes Chelan, Roosevelt, Pend Oreille, Koocanusa, and čłq̓etkʷ (Flathead). Lake Coeur d'Alene – bottom center – was the setting for the pilot study in this research



From its headwaters in the Rocky Mountains of British Columbia, Idaho, and Montana, the Columbia River's mainstem flows nearly 2,000 km before emptying into the Pacific Ocean along the border between Oregon and Washington. Various stems of the Columbia River system cross, and in some cases, re-cross the Canada-US border. Lakes Chelan, Franklin D. Roosevelt, Pend Oreille, and čłq̓etkʷ (Flathead) are entirely in three US states, Washington, Idaho, and Montana. Lake Koocanusa extends north from Montana into Canada, with nearly half of its length in British Columbia. Each reservoir in this study is created by or enhanced by a hydroelectric dam. Factors such as major land tenure, dam ownership, water user needs, and water quality and fisheries issues, however, vary between basins. Governance styles range from top-down co-management by dam owners and government agencies to fuller collaborations by a range of actors.

2.2. Obtaining network data

Our research used online surveys to collect data in the five Columbia River Basin reservoirs. Following a whole-network study approach, we defined the network boundaries to include all actors who participate in the water governance issues in the five basins, rather than by the basins' watershed delineation. For this research, water governance includes organizations involved in fisheries and water quality management. In our preliminary scoping we used resource management documents (Scheinert *et al.*, 2015), website searches, meeting minutes, to identify organizations that are likely to play a role in the basin (see Lubell & Lippert, 2011).

We first developed a name list of organizations for the five basins. This list included organizations involved in fisheries and water quality management. Study basins are located in different political units such as state or federal (e.g., part of Koocanusa basin falls in Canada). The name list of organizations was standardized by actor type, such as "state-DEQ", to encompass different state departments of environmental quality operating in the basins. We identified 59 actor types across the basins, including federal, tribe, state, regional, county, and city government agencies, educational and research institutions, irrigation or utilities providers, industry, and non-profits.

We reached out to each actor in each basin using a mixed-mode approach (Vaske, 2011) of email, phone, in-person contacts, and via public forums. Since management roles differ considerably across organizations, we asked for the person in each organization who was most involved in the management of either fisheries or water quality issues to fill the survey. Respondents also identified other actors that were not included in the original list, adding to it in a form of snowball sampling (Schneider *et al.*, 2003). The number of actor-type organizations varied by basin: 38 in Pend Oreille, 35 in Koocanusa, 32 in Flathead, 39 in Roosevelt, and 35 in Chelan.

A total of 75 surveys were filled of 179 across all five basins. An additional 18 actors asserted they have no active involvement in water governance, though some received numerous incoming ties. The response rate, by basin, was: Pend Oreille 51.3 percent, Koocanusa 48.7 percent, Flathead 61.8 percent, Roosevelt 46.3 percent, and Chelan 37.8 percent. Response rates were comparable to results in other water governance studies conducted in the U.S. and Canada (see Schneider *et al.*, 2003; Lubell & Lippert, 2011; Scheinert *et al.*, 2015; Horning *et al.*, 2017).

Surveys were organized into five thematic sections: (1) Information on the organization and the individual taking the survey; (2) Lake health indicators; (3) The network, or who interacts with whom; (4) Perceptions of network functions and success; (5) Whether or not, and how the public is engaged in the governance process. The current paper draws primarily on the network section of the survey, where respondents indicated whom they interact with and the purpose for contact, including seeking information or data; expertise or advice; technical collaboration; funding support; and political support.

2.3. Constructing and visualizing matrices

We used the contacts identified in surveys to construct socio-matrices of un-weighted, directed ties for all basins (actor A may reach out to B, but B may not reach back to A). For each basin, we created the combined water governance network matrix first, as many respondents marked contacts for both fisheries and water quality issues. Actors also indicated whether contacts were for fisheries, water quality, or both, allowing us to filter each basin matrix to a fisheries network and a water quality network.

Lake Koocanusa presented a unique challenge, as it straddles an international boundary. Many U.S. resource laws mandate collaborative co-management of resources[1]. Canadian laws are sparse in such requirements however, which is reflected in low diversity of central actors in governance networks (Horning *et al.*, 2017). To address these differences we created the Koocanusa matrix first from a Canada/ British Columbia (KBC) perspective,

[1] E.g.: The Pacific Northwest Electric Power Planning and Conservation Act of 2010, 16. U.S.C., Ch. 12H: 838(2) - 838(3)(B), *94 Stat. 2698*.

with the Canadian agencies coded as the primary, e.g. "federal" actors, and U.S. agencies as "federal-other". A second matrix reflected actor positions from a U.S./ Montana (KMT) perspective.

Non-responses create issues in network analysis, as they result in missing ties and errors in centrality measures. Imputing missing contacts using the median of 3-nearest neighbors technique, which is based on the smallest calculated (Euclidian) distance of incoming ties to actors, produces the most reliable results with up to 50 percent missing values (Žnidaršič *et al.*, 2017). To minimize the degree of imputation and to improve reliability, we reconstructed missing contacts using information from archival data and personal communications before imputing the remaining missing ties[2].

The survey showed that some basin networks were comprised largely of government agencies. Both the U.S. and Canada have strong environmental quality laws and standards that impact the Pacific Northwest[3], which are nevertheless administered quite differently at the national, or state and provincial levels. Tribes and First Nations, as sovereigns, operate their own natural resource departments. Regulatory structure coupled with centralized management suggests a dominance of tribal and government agencies in the centers of water governance networks. Thus, we first compare the patterns in each reservoir basin based on actor type – *who an actor is*.

2.4. Constructing the suite of roles

Our main premise, however, is that central actors in basin networks fulfill multiple distinct roles, conceived as a suite of roles – *what an actor does*. We identified three overarching types of roles: resource roles, formal roles, and informal roles, each constituting number of specific tasks or functions. We ascertained these roles and the corresponding tasks by using survey responses and planning documents, and by talking to practitioners in the study area. *Resource* roles are an aggregate of four tasks containing providing data and information, expertise, funding support, and/or political support. *Formal* roles are the sum of seven functions involving legal framework, including dam, land and resource tenure, regulatory oversight, regulatory obligation,

[2] Following recommendations of Žnidaršič and her co-author Patrick Doreian (separate personal communications).

[3] Environmental laws include: [U.S.] Clean Water Act (33 U.S.C §§ 1251 et seq.); Magnuson-Stevens Fishery Conservation and Management Act (16 U.S.C. §§ 1801 et seq.); Endangered Species Act (16 U.S.C. §§1531-1544); Canada Water Act (R.S.C., 1985, c. C-11); Fisheries Act (R.S.C. 1985, c F-14); and [Can.] Endangered Species Act (S.O. 2007, c 6).

treaty rights, formal agreements and court decisions. *Informal* roles of organizations are an aggregate of four functions including technical collaborations, hosting public forums for information exchange, publishing newsletters (print or email) or providing online information and resources, and leadership (Johnson *et al.*, 2003), or the degree to which an organization (or a key person within it) is considered by others to be a policy entrepreneur or "rock star" actor.

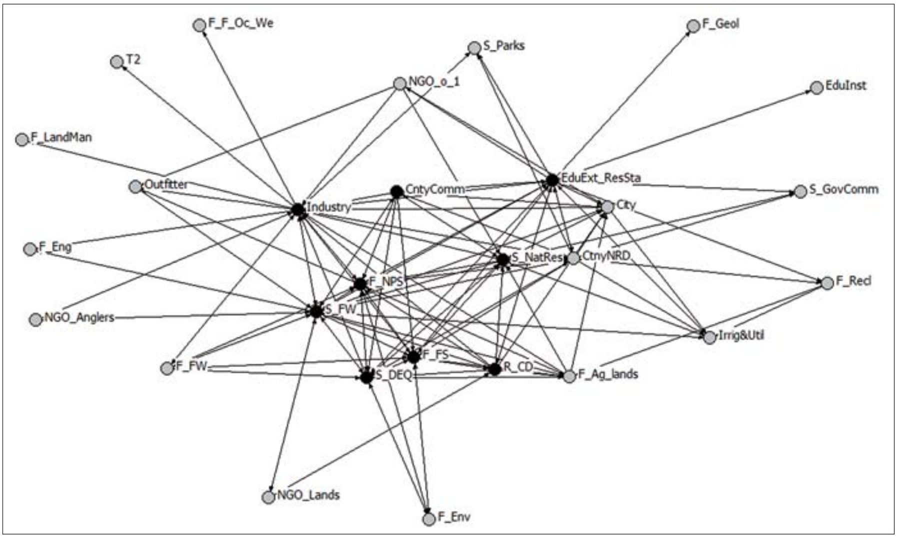
An actor can simultaneously serve in formal, resource, and informal roles. In Koocanusa, for example, the U.S. Army Corps of Engineers is the dam-owner and operator (a formal role), but it actively provides data, and funds research (resource roles), and actively collaborates on research with other actors, and co-hosts a well-publicized annual public forum (informal roles). We constructed actor by roles in which each organizational actor was assigned a 1 when it performed a function noted above, 0 otherwise. The sum of the 1s for each actor in the three role-type matrices resulted in a quantitative score measuring the level of each actor's engagement in the lake basins. The level of resource roles varied from 0 to 4, with an average of 2.8. The maximum number of formal roles was six (average of 0.6), and maximum four informal roles (average of 1.3).

3. Analysis

3.1. Network visualization

We visualized the network matrices using Ucinet for social network analysis (Borgatti *et al.*, 2002). Visualization of networks revealed differences in the densities and distributions of interactions both between and within basins, but all 18 matrices displayed core-periphery patterns. Figure 2 shows a combined fisheries and water quality network for Chelan basin, representing the classic core-periphery pattern observed across the basins. The core actors in the network, as blocked by Ucinet's core-periphery test, are densely tied to each other and are displayed in solid black whereas the periphery actors are shown in grey. It is also apparent that the core-periphery network is also dominated by a few central actors.

Fig. 2 – The combined fisheries and water quality network for the Chelan basin displays classic core-periphery network structure



3.2. Core-periphery tests

Core-periphery tests in UCINET were used to confirm the visual impressions of the 18 matrices. The core-periphery function in Ucinet uses a generic algorithm to divide actors into core and peripheral membership classes, and simultaneously fits a model to the network data, blocking the member classes into a 2x2 contact matrix. The observed networks are compared relative to an idealized core-periphery structure, producing one core-core region with a "perfect" 1-block (all connected), two core-periphery regions with 1-blocks, and one periphery-periphery region with a 0-block (no contacts) (Borgatti *et al.*, 2000). Core-periphery fit scores vary between 0 and 1. A resulting fit score that approaches 1 corresponds to a strong core-periphery structure.

Core-periphery tests in our study confirm visual observations with surprising consistency (Table 1). Chelan's network shows the highest core-periphery fit scores for combined network (0.823) and for fisheries (0.791) and water quality (0.750). Core-periphery fit scores are consistently on the high side in all basins, the minimum being 0.623 except for Pend Oreille, being 0.546.

Tab. 1 – Core-periphery fit scores of basin networks; a higher score indicates greater core-periphery structure. PO = Pend Oreille, KBC = Koocanusa from a Canadian perspective, KMT = Koocanusa from a U.S. perspective, F = Flathead, R = Roosevelt, C = Chelan

Network	*PO*	*KBC*	*KMT*	*F*	*R*	*C*
Full (Fish & WQ)	0.623	0.696	0.674	0.678	0.688	0.823
Fisheries	0.646	0.651	0.630	0.651	0.714	0.791
Water Quality	0.546	0.725	0.699	0.722	0.644	0.750

3.3. QAP correlation of network patterns on actor types

We used the quadratic assignment procedure (QAP) in R to assess correlations between network structures across the basins, as given by actor type. The technique has been used in evaluating similarities of patterns in watershed governance networks (see Scheinert *et al.*, 2015). Here we were interested in finding whether the same actors appear in the same structural positions between the various basins. QAP unstrings network matrices into vectors for pairwise matrix correlations. We used the algorithm's default of 1,000 permutations to determine statistical significance (Butts, 2008).

The QAP correlation results do not support similarity in network structure based on actor types. Among all the pairwise matrix correlations, the highest correlations found between Koocanusa, Canadian perspective (KBC) and Koocanusa U.S. perspective (KMT), which were 0.275 (fisheries) and 0.287 (water quality). Both network visualizations and Ucinet's core-periphery block models show different actors in the network cores.

3.4. Poisson regressions on the suite of roles

In testing the suite of roles argument, we are interested in how well resource roles, formal roles and informal roles explain high indegree centrality scores of core actors in the core-periphery structures of the five basin's water governance networks.

Prior to conducting regressions, we ran Pearson's correlation coefficient between the actor indegrees and the three different roles types, measured in counts (scores) as discussed earlier. Correlation values range from moderate to strong on combined networks (fisheries plus water quality networks), but are more variable in separate networks. The resource role correlates weakly with indegree in Flathead's fisheries network (0.396) but correlates well with Chelan's fisheries network (0.751). For the formal role, correlations varied from 0.513 with Pend Oreille's water quality network to 0.799 with Koocanusa's fisheries network. For informal roles, correlations are observed between 0.465 (Chelan's fisheries network) and 0.736 (Chelan's water quality network).

Next, we used Poisson regressions to examine effects of the suite of roles on centrality of actors, measured in indegrees. Poisson regression is the preferred statistical technique when the outcome variable such as ours is measured in counts (Coxe *et al.*, 2009, Legler & Roback, 2015). An actor's indegree, or degree centrality, captures the number of ties it receives (receiving a message or inquiry), reflecting its centrality vis-à-vis other actors in the network. We used actor's degree centrality as a proxy for being core or central actor in the network. We used the three types of roles – resource roles, formal roles, and informal roles – as the independent variables, measured as scores. In our case, high indegree (central) actors are often highly active communicators, thus also having high outdegrees. Since indegrees and outdegrees are different, yet correlated, phenomena, we included outdegree of actors as a control in our model, and expect significant association between the two. Multivariate specification allowed us to determine the effect of each role type on the centrality of actors controlling for other roles and outdegrees in the model.

Table 2 presents the regression results for effects of the suite of roles on indegree centrality on the aggregate of all five basins types on both fisheries and water quality networks. The reported coefficients are in exponentiated values with corresponding p-values. We also tested for potential multicollinearity of the independent variables but did not found to be an issue. Both models are significant as indicated by chi-square test. The models account for 44 percent proportional reduction in deviance due to the inclusion of the predictors as indicated by the McFadden pseudo R-square (see Coxe *et al.*, 2009). The estimated coefficients for all three role types in both fisheries and water quality networks are statistically significant, suggesting that the number of roles the organizations play determine their indegree centrality in these networks. For example, in the case of fisheries, an organization providing four resources is expected to have, on average, 1.53 times as many indegrees compared to an organization providing three resources, controlling for other roles and outdegrees. In other words, an organization is expected to increase its indegree centrality by 55 percent for each additional resource it provides. For an average organization having one indegree, a 53 percent increase with an additional resource role may not sound much. However, for a core actor with many indegrees, the 53 percent increase is substantial.

Tab. 2 – Poisson regression estimates explaining indegree centrality in aggregate across all basins. Coefficients are in exponentiated values; values in brackets are p-values; pseudo R-square is McFadden

Variable	*Fisheries all basins*	*Water quality all basins*
(Intercept)	0.69 (0.15)*	0.98 (0.13)
Resource	1.53 (0.05)***	1.44 (0.04)***
Formal	1.24 (0.03)***	1.24 (0.02)***
Informal	1.12 (0.04)**	1.19 (0.03)***
Outdeg	1.04 (0.01)***	1.03 (0)***
R-square	0.98	0.98
N	179	179

Sig. codes: 0.00 "***", 0.001 "**", 0.01"'*" 0.05

The importance of the suite of roles remains largely consistent across all reservoir basins, even when the results are separated by basin. The effects sizes and their significance levels vary by basin and network focus, however. In the fisheries governance networks (Table 3), the resource roles are important in all but Flathead basin. Formal roles affect indegree centrality in all networks except for Chelan. Informal roles are especially important in Flathead basin, but not in the other four. The outdegree has small effects on indegrees in Roosevelt and Chelan.

Tab. 3 – Poisson regression estimates explaining indegree centrality, by basin, for fisheries networks. Coefficients are in exponentiated values; values in brackets are p-values; pseudo R-square is McFadden

Variable	*Pend Oreille*	*Koocanusa*	*Flathead*	*Roosevelt*	*Chelan*
(Intercept)	0.43 (0.33)*	1.00 (0.29)	1.18 (0.62)	1.27 (0.31)	0.26 (0.47)*
Resource	1.68 (0.1)***	1.49 (0.09)***	1.3 (0.18)	1.39 (0.09)***	1.59 (0.16)***
Formal	1.49 (0.08)***	1.28 (0.1)*	1.17 (0.05)***	1.24 (0.05)***	1.69 (0.24)
Informal	1.17 (0.09)	1.15 (0.09)	1.28 (0.08)**	1.00 (0.1)	0.74 (0.21)
Outdeg	1.03 (0.01)	1.02 (0.02)	1.02 (0.01)	1.04 (0.01)***	1.18 (0.05)***
R-square	0.99	0.98	0.88	0.98	0.92
N	38	35	32	39	35

Sig. Codes: 0.00 "***", 0.001 "**", 0.01 "*" 0.05

In the water quality governance networks (Table 4), resource roles affect the indegree in all networks. Formal roles, again, are important in all basins except Chelan. Informal roles are highly variable in their impact on indegrees, having effects only in the Pend Oreille and Roosevelt basins. We removed informal roles from the model for the Flathead water quality network due to collinearity between informal and formal roles. Outdegree again has small, but significant effects.

Tab. 4 – Poisson regression estimates explaining indegree centrality, by basin, for water quality networks. Coefficients are in exponentiated values; values in brackets are p-values; pseudo R-square is McFadden

Variable	*Pend Oreille*	*Koocanusa*	*Flathead*	*Roosevelt*	*Chelan*
(Intercept)	0.9 (0.25)	1.18 (0.29)	2.13 (0.47)	1.02 (0.31)	0.48 (0.33)*
Resource	1.52 (0.08)***	1.36 (0.09)***	1.28 (0.13)*	1.41 (0.09)***	1.49 (0.11)***
Formal	1.25 (0.07)***	1.27 (0.09)*	1.24 (0.04)***	1.25 (0.05)***	1.3 (0.16)
Informal	1.36 (0.08)***	1.13 (0.08)	†	1.22 (0.09)*	1.08 (0.1)
Outdeg	1.01 (0.01)	1.05 (0.02)**	1.04 (0.01)***	1.03 (0.01)**	1.1 (0.03)***
R-square	0.97	0.98	0.90	0.97	0.97
N	38	35	32	39	35

Sig. Codes: 0.00 "*** ", 0.001 "** ", 0.01 "* " 0.05

† variable dropped due to multicollinearity issues

4. Discussion

Both network visualization and core-periphery analyses found the presence of core-periphery governance network structure in all five basins. It is notable that no single actor is central, rather a group of core actors, is in the apex of the basin governance. The core actors are diverse organizations acting collaboratively for collective wellbeing, reflecting a Ostrom's (2010) polycentric system of governance. Densely connected core groups create a condition for the core actors to regularly consult when making decisions. The mechanism also allows them to build social capital as information flows quickly through overlapping ties among the core actors. Links between the peripheral and core actors create governance systems that provide opportunities for peripheral actors not only to participate but also to voice their interests or concerns via core actors.

As to what drives basin actors to be at the core of the network, the very low correlation observed between structurally similar core-periphery networks across all basins shows that the actor type has little influence. Since all five basins have core-periphery network structures and 27 actors are consistently present in at least four of five basins – ten of which are federal agencies – significant actor types defining networks of ties would be reflected in high correlation across the basin networks. Lacking similarities in actor-positions in the networks, however, suggest that another driver is far more important to network centrality scores than who an actor is.

Analysis results support our premise that the emergence of core actors in the networks of the study basins is driven by filling multiple roles – what an actor does. The generally high correlations between the three types of roles and indegree centrality of actors suggest that, when making the choice of partners, actors care deeply what those partners have to offer to meet the governing challenges of the basin. As we expected, organizations in the study basins tend to rely on other organizations that play multiple roles, have greater capacity, or can shoulder more responsibility. This mechanism could lead to highly centralized networks with a dominant central actor, but in this study, basin networks are characterized by the presence of multiple central (core) actors. This suggests that actors in the basins favor groups of central actors that encourage greater consultation and collaboration in the network.

5. Conclusion

Our research shows that lake basin governance in the Columbia River Basin is characterized by core-periphery network structure, and the centrality of actors in the network is driven by what actors do rather than what actors are. In this research, we examined three types of roles – resource roles, formal roles, and informal roles. We used social network analysis and conventional statistical analysis in a stepwise manner to examine our proposition that roles matter. The significance of specific types of roles varies across the basins. While more study is needed for better understanding, part of the variability appears to be associated with the size and composition of the core actors and how roles are distributed across these actors. Variability in actor composition and network evolution appears to stem also from geographic, cultural, and historical contexts of each basin. Further, even within the fisheries or water quality networks, multiple sub-issues are specifically addressed by smaller sets of specialized actors.

Core-periphery structure in collaborative water governance is not uncommon; yet, the literature is scant, or limited to a single basin. This research contributes to this literature by highlighting the presence and importance of core-periphery network structures in multiple lake basins. We were able to observe the network governance structure in the basins at one point in time. It is possible that the governance structure changes with time, requiring the study of basin networks over time. In addition, five basins still are a small number study limiting the generalizability of the results. Future research should address these limitations. Applying this study to additional basins may reveal similarities in network characteristics of central actors, such as organization size, or some other features that are and distinctly different between and among the basins.

References

Bodin Ö., Ramirez-Sanchez S., Ernstson H., Prell C. (2011), "Introduction. A social relational approach to natural resource governance", in Ö. Bodin, C. Prell (*eds.*), *Social Networks and Natural Resource Management: Uncovering the Social Fabric of Environmental Governance*, Cambridge University Press, Cambridge, pp. 3-28.

Borgatti S.P., Everett M.G. (2000), "Models of core/periphery structures", *Social Networks*, 21 (4), pp. 375-395.

Borgatti S.P., Everett M.G., Freeman L.C. (2002), *Ucinet for Windows: Software for Social Network Analysis*, Analytic Technologies, Harvard (MA).

Butts C.T. (2008), "Social network analysis with sna", *Journal of Statistical Software*, 24 (6), pp. 1-51.

Coxe S., West S.G., Aiken L.S. (2009), "The analysis of count data: A gentle introduction to Poisson regression and its alternatives", *Journal of Personality Assessment*, 91 (2), pp. 121-136.

Csermely P., London A., Wu L.Y., Uzzi B. (2013), "Structure and dynamics of core/periphery networks", *Journal of Complex Networks*, 1 (2), pp. 93-123.

Folke C., Hahn T., Olsson P., Norberg J. (2005), "Adaptive governance of social-ecological systems", *Annu. Rev. Environ. Resour.*, 30, pp. 441-473.

Haunss S., Schmidtke H., Biegoń D. (2017), "Integriertes multidimensionales Design", *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69 (2), pp. 185-212.

Horning D.G., Bauer B.O., Cohen S.J. (2017), "Watershed governance for rural communities: aligning network structure with stakeholder vision", *Journal of Rural and Community Development*, 11 (2), pp. 45-71.

Jasny L., Lubell M. (2015), "Two-mode brokerage in policy networks", *Social Networks*, 41, pp. 36-47.
Johnson J.C., Boster J.S., Palinkas L.A. (2003), "Social roles and the evolution of networks in extreme and isolated environments", *Journal of Mathematical Sociology*, 27 (2-3), pp. 89-121.
Kapucu N., Hu Q., Khosa S. (2017), "The state of network research in public administration", *Administration & Society*, 49 (8), pp. 1087-1120.
Krebs V., Holley J. (2006), *Building smart communities through network weaving. Appalachian Center for Economic Networks*, self published.
Legler J., Roback P. (2015), *Broaden your statistical horizons: generalized linear models and multilevel models*, https://bookdown.org/roback/bookdown-bysh/.
Lubell M., Lippert L. (2011), "Integrated regional water management: a study of collaboration or water politics-as-usual in California, USA", *International Review of Administrative Sciences*, 77 (1), pp. 76-100.
Onwuegbuzie A.J., Collins K.M. (2007), "A typology of mixed methods sampling designs in social science research", *The Qualitative Report*, 12 (2), pp. 281-316.
Ostrom E. (2010), "Polycentric systems for coping with collective action and global environmental change", *Global Environmental Change*, 20 (4), pp. 550-557.
Pfeffer J., Salancik G.R. (2003), *The external control of organizations: A resource dependence perspective*, Stanford University Press, Redwood City (CA).
Scheinert S., Koliba C., Hurley S., Coleman S., Zia A. (2015), "The Shape of Watershed Governance: Locating the Boundaries of Multiplex Networks", *Complexity, Governance & Networks*, 2 (1), pp. 65-82.
Schneider M., Scholz J., Lubell M., Mindruta D., Edwardsen M. (2003), "Building consensual institutions: networks and the National Estuary Program", *American Journal of Political Science*, 47 (1), pp. 143-158.
Scholz J.T., Stiftel B. (*eds.*) (2010), *Adaptive governance and water conflict: new institutions for collaborative planning*, Routledge, London.
Smedstad J., Gosnell H. (2013), "Do adaptive comanagement processes lead to adaptive comanagement outcomes? A multicase study of long-term outcomes associated with the National Riparian Service Team's place-based riparian assistance", *Ecology and Society*, 18 (4).
Vaske J.J. (2011), "Advantages and disadvantages of internet surveys: Introduction to the special issue", *Human Dimensions of Wildlife*, 16 (3), pp. 149-153.
Wasserman S., Faust K. (2009), *Social network analysis: Methods and applications*, Vol. 8, Cambridge University Press, Cambridge.
Yang J., Leskovec J. (2014), "Overlapping communities explain core-periphery organization of networks", *Proceedings of the IEEE*, 102 (12), pp. 1892-1902.
Žnidaršič A., Ferligoj A., Doreian P. (2017), "Actor non-response in valued social networks: The impact of different non-response treatments on the stability of blockmodels", *Social Networks*, 48, pp. 46-56.

7. Development program intervention and network building. Application of bipartite Exponential Random Graph Models

by Manoj K. Shrestha*

1. Introduction

Most social network research assumes that network building behavior of actors is influenced by the potential consequence of the choice of partners. Networks are generally self-organized as actors choose to connect based on the utility of contact (Ostrom, 2009). However, the consequence-based choice overlooks the importance of social interventions aimed at enabling actors, affecting the choice of partners. Actors face barriers to creating ties. While some actors may be introvert, others may lack conducive environment or skills needed to engage socially (Chaskin *et al.*, 2001). Social interventions are activities designed to improve social skills of actors or to modify the social context within which they function. These interventions can take various forms, including creating formal or informal forums to socialize and interact and providing trainings on social engagement and communication skills, the goal being improved network building capacity of actors.

Social interventions, or capacity building measures, are common in community development assistance programs. Many community development programs invest in capacity building measures as part of the assistance to communities. The premise is that the communities with exposure to social interventions will engage in network building not only to complete the assisted project but also to pursue self-help activities after the assistance ends. Programs that produce effects far beyond the scope of the project itself is deemed the most successful development assistance (Goldin, Rogers & Stern, 2010). However, there is little or no study of whether and how exposure to social interventions influence network building in public assistance programs.

* Department of Politics and Philosophy, University of Idaho, mks@uidaho.edu.

The purpose of this research is to test whether social interventions enable the aided communities to build networks in the long-term after the assistance has ceased. This research builds on earlier works that examined the effect of communities' network with organizations on the communities' ability to secure program funds during the project implementation (Shrestha, 2013) and communities' network activities in the short-term after the program funding is over (Shrestha, 2019). It uses comparison group evaluation to assess the network activity of rural communities in Nepal who received aid through the community water supply program. This research contributes to the knowledge of the process theory of public assistance programs by evaluating the assumptions underlying the network building behavior. Given that networks are ubiquitous, determining the effect of social intervention on network building is especially important. The resulting insight can be valuable to management looking into modifying network behavior for the desired result.

This research proceeds with the discussion of a process theory of network building, explaining how social interventions affect network building. Next, it outlines the research design, detailing collection of network data, design of comparison group evaluation and specification of bipartite exponential random graph models to model the observed network. Then, it presents and discusses the results. It concludes with the implications and limitations of the research.

2. Social Network Intervention and Network Building

Networks are formed as actors engage in network building – an act of creating ties with other actors. Creating ties involves choosing partners to create ties with. The choice of a partner is influenced by the qualities it possesses or the positions it holds because of the presence of other ties (Robins *et al.*, 2007). For example, a focal actor may choose a partner to access resources it holds, or it may choose a central actor, preferred by many others, to achieve coordination or to avoid duplication (Provan & Milward, 1995). Besides, the focal actors' choice of partners can also be affected by the social skills they possess and the social environment they encounter. Actors differ in their social skills and in their ability to act in a certain social context. Social skills involve communication skills, interaction skills, conflict resolution, participation, and treating others respectfully (Chaskin *et al.*, 2001). Social environment includes the social setting, social institutions, power relations, and cultural norms and values that affect social interaction. Knowing how to initiate, respond and react, how to talk and gesture, and how to cope can significantly improve personal and professional relationships (Chaskin *et al.*, 2001). Understanding the subject matter, processes, and the roles the potential partners play can prove important in building relationships. (Argyle, Furnham & Graham, 1981).

Community capacity building literature stresses the existence of collaboration, problem-solving abilities, planning and decision-making skills and relationship building as important elements of community capacity (Chaskin *et al.*, 2001). Many community programs invest in community capacity building through funds set aside for technical assistance, separate from capital assistance meant for building physical assets. Although designed as short-term interventions, these measures are argued to build the long-term capacity of communities (Chaskin *et al.*, 2001). Technical assistance to communities commonly focuses on providing formal and on-the-job trainings. Formal training includes communication skills, conflict resolution, community engagement, record keeping, and collaboration and relationship building. On-the-job training includes participation in planning, implementation and operation of the funded projects, organizing community meetings, keeping minutes, and development of community institution (e.g., Users' Committee) to sustain the project long after its completion.

Community development programs assume that the capacity building measures will improve communities' capacity for network building which, in turn, will lead to greater network activity of the communities not only to access resources for the assisted project but also to mobilize support for new self-help projects after the assistance is over. Increased network activity can occur when communities are motivated to do so. Communities hardly possess all the resources they need to successfully plan and implement projects. According to the resource dependence theory, communities will reach out to organizations to access resources to achieve their goals (e.g., Pfeffer & Salancik, 1978). Citing the example of rotating savings and credits associations, Granovettor (1985) argues that while members of these group initially draw upon the resources of family and peers but then attempt to forge broader and more autonomous ties beyond the group as their need for larger markets and more sophisticated inputs expand. In the context of local economic development, Woolcock (1998) suggests that a community's internal social capital serves as a basis for launching development initiatives, but it must be complemented over time by forming linkages to broader extra-community institutions to get ahead. Studies carried out in the US, Canada, and developing countries support this theory (e.g., Agnitsch, Flora & Ryan, 2006; Dale & Newman, 2008; Woolcock & Narayan, 2000).

Most research in this field is limited to the effect of network building on outcomes during the program implementation. While the attributes of contacted organizations such as types of organizations (see, Shrestha, 2019) or the composition of contacted organizations are important in affecting outcomes, the role of network activity in general has been a subject of constant interest in the literature on organizational networks. For example, O'Toole and Meier (2004) found that school superintendents' frequency of contacts with school boards, business leaders, state legislators, education agency, and other superintendents were correlated with the performance of schools. Shrestha (2013) found that communities' organizational network size (i.e., degree centrality) was associated with their success in securing the program funds needed to implement water supply projects. There is limited research on the effect of network activity after the implementation of a program or the program funding is over, especially when it relates to longer-term network activity. In addition, there is limited knowledge on the effect of program exposure on network building. Shrestha (2019) examined post-project network building of rural communities with various organizations but it was limited to short-term network activity. Besides, the communities in the study were limited to those that received program funding.

This research attempts to fill these two gaps in the literature. It examines communities' long-term network activity, which is one of the important network building behaviors for outcomes to realize. It also probs whether the communities exposed to program intervention display greater network activity than the communities without the program exposure. Because networks are ubiquitous, communities, with or without program interventions, are likely to engage in networking. The expectation is that the communities with program exposure will show greater network activity than the communities without such an exposure.

3. Data, Measures, and Analysis

3.1. The empirical context

The Rural Water Supply and Sanitation Program (RWSSP) in Nepal provided the empirical context for this research. Supported by the World Bank, the immediate goal of the program was to help rural communities to improve their access to clean drinking water. The program adopted a collaborative model. Each interested community initiated a water supply project and elected a Water Supply and Sanitation User Committee (WSUC) to perform the day-to-day tasks of the project. The program, for its part, supported the communities with technical assistance and matching funds to implement the water projects. The program used criteria-based funding. The criteria included securing the matching cash and in-kind (mainly unskilled labor) portions of the project's costs, complying with the regulatory provisions, and ensuring transparency by conducing open public meeting about the operation and expenses of the project. In addition, communities that were from remote areas, smaller in size, and were to save more time in collecting water with new project received priority for the program funds.

Altogether 125 communities applied for funds from the program. Over half of them were funded. As part of the technical assistance of the program during the implementation of the project, the funded ones received on-site formal and on-the-job training as well as advise relating to the importance of linkages with various organizations to mobilize various support they need to implement their projects. The capacity building measures together with criteria-based program funding encouraged the communities to create ties with different organizations including governmental, nonprofits, and business organizations to mobilize their support to build the water projects. Such support included the community's portion of matching funds, regulatory support, expertise, or political support. The program also hoped that the communities will use the improved capacity to connect with organizations to mobilize resources needed to address other community needs after the funded water project is complete, which is the focus of this research.

3.2. Obtaining community by organization network

Communities' ties, representing resource contacts, with organizations were collected from the field survey of 125 communities, spread out in five districts of the Central Development Region of Nepal. These communities applied for the program funds for water projects and, thus, included the communities that received the program funds (and the capacity building interventions) and the communities that did not[1]. Districts are the main sub-national political and administrative hubs in the country. Their boarders include village and district local governments, local offices of governmental agencies, nonprofits, businesses, and community organizations. The contacts were aimed at mobilizing resources to purse self-help activities. The survey was carried out in 2014 from February through April after 8 years since the funding decision was made in 2006; that is 7 years after the completion of the water projects by the funded communities.

[1] The districts are Kavre, Sindhupalchowk, Dhading, Makwanpur, and Sindhuli.

The filed survey used a pre-designed questionnaire. The locally hired interviewers conducted the survey by visiting the communities in their villages. For each community interviewed, the respondents were the WSUC as a group which included the Chairperson, the Secretary, and the Treasurer. The respondents were asked to name the organizations they contacted to mobilize resources needed to pursue self-help projects. The survey asked for organizational contacts made after 2007, the year of completion of the water projects by the funded communities, for information, funds, political buy-in, and regulatory support, deemed important for pursuing community projects. The respondents were asked to name the organizations they contacted to mobilized resources for those projects. Four open-ended name-generator questions were asked. The opening question was: "Contacts with organizations can occur to address key resource challenges (information, fund, political support, and regulatory help) that you encounter. Now, please recall and tell us all the governmental, non-governmental, and private organizations; political parties; local governments; community organizations; and other stakeholders at the local, district and central levels with whom you made one or more contacts". The remaining questions were repeats of the opening question which were asked separately to capture the organizational contacts made at the local, district, and central levels. The face-to-face interview in the local language, Nepali, combined with judicious use of prompts improved the reliability of the responses. This resulted a reasonably complete contact data. The survey also collected data on community size and the communities' opinion about conflict within community and future collaboration.

Together, 125 communities contacted 478 organizations, resulting a bipartite (two-mode) network of size 125x478. A bipartite network consists of two sets of actors (e.g., communities and organizations) and ties are defined between pairs of actors belonging to different sets (Wasserman and Faust, 1994). These communities contacted a wide variety of organizations including government agencies, local governments, banks, cooperatives, nonprofits, local schools, and political parties, for example. The spinoff projects pursued included village roads, irrigation channels, school buildings, trails, bridges and community centers.

Table 1 presents the summary statistics of the organizations contacted by the communities and their attributes with a breakdown for communities that received program funding and the communities that did not.

The community size or the number of households, varied from 26 to 409 households. The level of within-community conflict and perceived future collaboration with organizations were measured in a Likert scale of 1 through 7 where 1 indicates no within-community conflict or no future collaboration and 7 indicates excessive conflict or maximum future collaboration. On average, the communities display low level of conflict and greater willingness to collaborate with organizations.

Tab. 1 – Summary statistics of communities contacting organizations

Variables	*Mean*	*Median*	*StDev*	*Min.*	*Max.*
Funded communities					
No. of org. contacted	6.49	6	4.76	1	19
No. of outside org. contacted	3.47	3	1.70	0	9
No. of households (size)	131.52	130	50.14	30	280
Community conflict	2.12	2	1.44	1	6
Future collaboration	5.09	5	1.52	3	7
Unfunded communities					
No. of org. contacted	6.27	5	5.17	2	18
No. of outside org. contacted	3.11	3	2.31	0	10
No. of households (size)	109.69	92.5	76.06	26	409
Community conflict	1.98	1.5	1.38	1	6
Future collaboration	4.84	4	1.79	2	7

Note: Data from field survey. Community conflict and future collaboration are in Likert scale of 1 through 7.

3.3. Evaluation design

3.3.1. Comparison group evaluation design

The first step involved comparing the average network activity between the communities that secured program funding and the communities that did not. To assess the effect of program funding on network activity, the posttest-only comparison group evaluation design is used. The comparison group design uses a nonrandomly assigned comparison group with the treatment group unlike in a control group design which uses a randomly assigned control group for comparison (Rossi *et al.*, 2019). Two groups of communities with and without the program funding provided the setting for this evaluation. Out of 125 communities applied for the program funds, the treatment group consisted 66 communities that received the program funds and the comparison group consisted 59 communities that did not receive the program funds. In this design, the effect of program exposure on network activity is the difference between the average level of network activity for the communities in the treatment group and in the comparison group.

The communities in both groups were from the same study areas and, thus, were expected to be similar geographically, culturally, and demographically. In addition, community size, within-community conflict, and willingness for future collaboration for both groups were compared to ensure that groups were similar except for the capacity building measures that came with program funding. The combination of the covariates and the selection of the comparison group from the same geographical or cultural settings is expected to reduce the bias in the estimates of the effect of program intervention (Rossi *et al.*, 2019).

The average community size, or the number of households, for the treatment group (i.e., funded communities) was 131.52 and for the comparison group (i.e., unfunded communities) was 109.69. In general, the larger communities face greater needs which means pursing more new projects of different types leading to seek ties with many organizations to mobilize a greater variety of expertise and resources than that of the smaller communities. The average level of within-community conflict for the treatment group was 2.12 and for the comparison group was 1.98 (in a Likert scale of 1 to 7), indicating negligible conflict within communities in both groups. Similarly, the treatment group's average level of willingness for future collaboration with organizations was 5.09 whereas, for the comparison group, it was 4.84 (in a Likert scale of 1 to 7), indicating that both groups of communities were moderately optimistic about future collaboration.

Network activity of a community was measured by its degree centrality. Degree centrality is the measure of an actor's number of ties (Wasserman & Faust, 1994). A community's degree centrality is the number of organizations contacted by the community. In the two-mode network matrix between communities listed in rows and organizations listed in columns, a community's degree centrality is the sum of all the 1s, or the presence of ties with organizations, over all the cells across a row, representing the community. A community with high degree centrality means that it has contacted greater number of organizations than a community with low degree centrality. The median degree centrality for the funded communities was 6 and for the unfunded communities was 5. Because network activity and covariates were measured at the same time for both groups of communities, internal validity is expected to be higher (see Rossi *et al.*, 2019).

3.3.2. Bipartite exponential random graph model

Ties are not independent and communities' degree centrality with organizations is expected to vary across the communities. Generally, the larger communities have greater community development needs and, therefore, likely to purse greater number and types of new projects than the small communities, leading to greater variance in the distribution of community degree centrality. Since a simple difference of means between two groups of communities does not take into account the structure of ties, a bipartite exponential random graph model (ERGM) was used to model (account for) the tendency of community activity spread (discussed below) along with the community attribute "funded community" to determine whether the probability of creating a tie is higher for the funded communities than for the unfunded communities. An ERGM is appropriate for statistical analysis of social networks observed cross-sectionally. Networks are patterns of local configuration of ties. An ERGM is a probability distribution reflecting our assumptions on the mechanisms that might have generated the observed network. The model describes the probability of a tie conditional on the rest of the network and the attributes of actors (Lusher, Koskinen & Robins, 2013; Wang, Pattison & Robins, 2013)[2]. Thus, with both structural and attribute effects included, these models allow to determine whether an actor with a particular attribute (in this case, the funded communities) has higher probability of creating ties than without the attribute, taking into account the network structures and other attributes. In terms of conventional multivariate analogy, in ERGMs, the observed matrix of ties is treated as the dependent variable and assumed structural features and node attributes are treated as the independent variables. ERGMs are the probability distribution from which the random graphs are simulated using the Markov Chain Monte Carlo (MCMC) methods (Lusher *et al.*, 2013). The simulations are then used to refine the initial estimation of the parameters (e.g., community activity spread) by comparing the value of the statistics computed on the simulated networks with those observed. The parameters are determined using the maximum likelihood estimation that maximize the probability of the observed network under the specified ERGM. The estimated parameter indicates whether a local configuration is more or less likely in the observed network than in a network where ties are created at random (Robins *et al.*, 2007). The estimated parameters are interpreted as logit coefficients, or conditional log odds, in a logistic regression (Robins *et al.*, 2007).

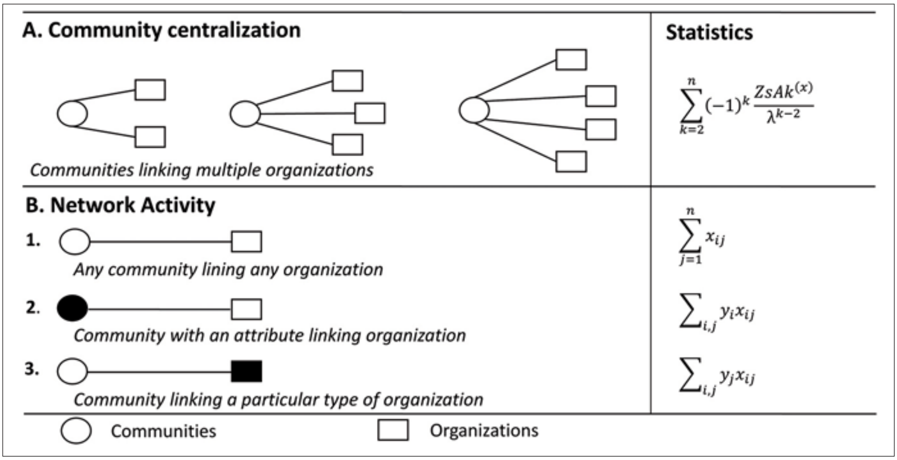
[2] For details on the general form of Exponential Radom Graph Models, see Lusher, Koskinen, and Robins, 2013; Wang, 2013.

In this model, the effect of funded communities on network activity is determined by including a dummy variable, coded 1 for the funded communities and 0 for the unfunded communities. This effect is captured by the "sender effect" (see Wang, 2013) as shown in figure 1B2, in which funded communities are represented in dark circle and organizations are represented in light squares. A positive parameter for the actor attribute effect for program funding means that the communities with program exposure have higher probability of creating network of ties with organizations than the communities without the program exposure.

The model also includes community degree centralization and several other attributes of the communities. Community degree centralization, also called community "activity spread" (see Wang, 2013), is the tendency of a community to connect with many organizations to mobilize wider support. The larger communities face greater needs than the smaller communities. This means that larger communities are likely to purse more new projects of different types leading to create ties with a broad set of organizations to mobilize a greater variety of expertise and resources. This tendency leads to a few communities more central in the network, suggesting a greater variance in the communities' degrees with organizations. Figure 1A shows 2-, 3-, and 4-star configurations centered on communities which mirror degree distribution of the communities. In ERGMs, degree distribution across the actors is parameterized by combining them into one parameter, called "alternating k-out-star" (Wang, 2013). The alternating *k*-star statistics is a weighted sum of all the stars in the network. A k in the alternating k-out-star parameter represents the size of star configuration (e.g., $k = 2, \ldots, 4$ in Figure 1A). A positive parameter for alternative *k*-out-stars suggests that the network activity is centralized around a few high-degree communities whereas a negative parameter means a more uniform distribution of the communities' degrees with organizations.

Community size, community conflict, and future collaboration were included to adjust likely differences between the funded and unfunded communities. The communities with higher values on those attributes are more likely to connect with organizations, which is captured by "sender effect" (see Wang, 2013). Figure 1B2 shows this effect in which the communities with these attributes are shown in dark circle. Moreover, the communities requiring more resources are more like to connect with organizations beyond their villages. This pattern is captured by Figure 1B3 in which outside organizations are represented by a black square. To capture this effect, the contacted organizations were categorized into outside and village organizations, and then a dummy variable was included, coded 1 for outside organizations and 0 for the village organizations. The outside organization variable is an aggregate of the district and the central level organizations. In addition, four study district dummy variables were included to adjust any other unobserved differences between the funded and unfunded communities. The statistics for the structural and attribute parameters are included in Figure 1.

Fig. 1 – Two-mode network configurations



Source: Adapted from Shrestha (2019)

4. Results and Discussion

Table 2 shows the average characteristics of the funded and unfunded communities. The average characteristics of the communities in both groups are similar, except for the average community size for the two groups of communities. Even for community size, the difference of means test is statistically significant at the margin. The average difference is 22 households, or 20% variation in the average size of the communities. The mean difference is small for within-community conflict and future collaboration.

Tab. 2 – Balance between funded and unfunded communities

Community characteristics	*Funded communities*	*Unfunded communities*	*Diff.*	*p-value**
Community size	131.52 (130)	109.69 (93)	21.83	0.06
Community conflict	2.12 (2)	1.98 (1.5)	0.14	0.54
Future collaboration	5.09 (5)	4.84 (4)	0.25	0.32

* Independent-samples t-test. Figures in the parentheses are median values.

Table 3 reports the average degree of the funded and unfunded communities with organizations. While the impact of the program exposure on the communities' degree is positive, the difference is small or the percentage variation in the average degree of the communities is 3.5%. The difference of means test suggests that the difference in the average degrees between the communities in the funded and unfunded groups is not statistically significant.

Tab. 3 – Comparing means of community degree with organizations

	Funded communities	*Unfunded communities*	*Diff.*	*p-value**
Community degree	6.49 (6)	6.27 (5)	0.22	0.28

* Mann-Whitney test. Figures in the parentheses are median values.

Table 4 reports the bipartite ERGM estimates for both structural and attributes parameters included in the model. The MPNet program available at http://www.melnet.org.au/pnet/ for analyzing bipartite networks was used to obtain the estimates (Wang *et al.*, 2014)[3]. The table also includes observed graph statistics and t-ratio indicating that the estimated model is well converged[4]. The results show that the funded-community attribute parameter is not statistically significant. This suggests that the funded communities are not different in the probability of creating ties with organizations than the unfunded communities controlling for community activity spread and other attributes in the model. The community activity spread parameter was statistically significant without community attributes but when the attributes were included the activity spread tendency dissipated. In general, communities are more likely to create ties with organizations outside their villages, indicated by the significant parameter for this effect. In general, the central or the district level organizations hold more information, resources, and authority and, thus, are likely to offer the communities more resources than the village level organizations. This finding is important and consistent with the economic development literature that suggest that extra-community ties are essential for communities aiming to get ahead (Woolcock, 1998). The parameter for community conflict is significant, indicating that the communities with greater internal conflict are

[3] The MPNet program labels for the configurations in Figure 1 are as follows: Community activity spread (centralization) = XASA; organization type (receiver effect) = XEdgeB; community attribute (sender effect) = XEdgeA; Edge parameter (any community linking with any organization) = XEdge (see, Wang *et al.*, 2014).

[4] Estimates were attempted with 2-stars, 3-stars, 3-paths, and 4-cycles and their different combinations; however, the models did not converge.

more likely to create tie with organizations. Given that communities prefer contacts with outside organizations, the communities may be compensating fallouts for internal conflict by seeking resources from outside. Communities with positive attitude about collaboration have high probability of creating ties with organizations. The community size is not significant in influencing ties with organizations; however, difference in sample district tend to exist affecting the communities' likelihood of creating ties with organizations.

Tab. 4 – Bipartite ERGM estimates for communities' probability of ties with organizations

Parameter	*Count@*	*Est (se)*	*Est (se)*	*t-ratio*
Edge	785	-5.67 (0.35)*	-5.71 (0.52)*	-0.03
Community activity	1,109	0.71 (0.19)*	0.06 (0.29)	-0.03
Funded community	422	0.03 (.07)	-0.07 (0.07)	-0.04
Outside orgs	452		0.77 (0.07)*	-0.02
Community size	95,746		0.001 (0.001)	0.01
Community conflict	1,887		0.12 (0.03)*	0.02
Future collaboration	4,090		0.15 (0.03)*	-0.01
Sindhupalchowk	169		0.10 (0.12)	-0.07
Makwanpur district	124		-0.45 (0.12)*	0.08
Sindhuli district	109		0.11 (0.13)	-0.01
Dhading district	87		-0.60 (0.15)*	0.02

* $p < 0.05$. @ Counts are observed graph statistics. The estimated models are well converged and t-ratios for convergence are between 0.01 to 0.08. The GOF t-statistic is < 0.01, meaning the fitted model compares well with the observed network. Outside (vs. village) organization measures "receiver effect"; all other variables in the model measure "sender effect"

The result for the funded community attribute obtained from the ERGM estimate that included structural and community attribute parameters was found similar to the result obtained from bivariate analysis. In both instances, the funded communities' probability of creating ties with organizations in the long-term – after 8 years of program funding – is not different from the unfunded communities. This finding is at odds with the finding of an earlier study that showed that, during the implementation phase of water projects, the funded communities were associated with greater network activity (Shrestha, 2013). The differences in network activity between funded and unfunded communities might fade over time because both groups of communities can overlap in their contacts with organizations, creating an opportunity for the unfunded communities to learn the strategies, used by the funded communities, through the shared organization. For example, one such organization in each study district is the District Development Committee (DDC) office – a go-to place for both the funded and unfunded communities for funding and other referrals. Likewise, although these communities are geographically scattered, it is possible for members of both groups of communities to meet at some common public venues such as a local market, allowing the unfunded communities to get information about the potential organizations to seek contacts with. Another plausible explanation could be that these communities, being located in remote areas, have low level of community capacity and, thus, require longer program support. In that case, when the program assistance ends, the capacity of funded communities, needed to actively engage with organizations, may recede over time. Finally, during this period, the unfunded communities may have secured external program support to improve their capacity. In this situation, the unfunded communities are likely to catch up with the capacity of the funded communities in building networks with organizations.

5. Conclusion

This research used comparison group evaluation to discover whether external program support increases the network activity of receiving communities in the long-term. Difference of means test as well as bipartite ERGM results show that the funded communities' likelihood of creating ties with organizations is not different from the unfunded communities. The differences in network activity between funded and unfunded communities can dissipate over time because the unfunded communities can learn from the funded communities either through the shared organizations or through participation in common venues such as local markets. It is also possible for the unfunded communities to secure external program support to improve their capacity to catch up with the funded communities. Nevertheless, these communities tend to create ties with outside organizations as these organizations hold greater resources than the village level organizations. Optimism about future collaboration also appears to contribute to network building behavior of these communities. While additional research will be needed to be conclusive about the absence of sustained impact of program intervention on network building, this finding, however, raises questions whether the extent and duration of the program support received by communities are adequate to make a difference in network activity. This is more likely for communities located in remote areas whose capacity is generally low, requiring longer support to sustain the ability to engage in network building.

From a theoretical perspective, this research highlights the need for examining not only the link between the extent and duration of external support and network building but also why network activity of actors fades over time. From the methodological perspective, this research advances the use of comparison group evaluation methods in social network analysis to assess the impact of social interventions on network building. The methodology used in this research should help start developing network evaluation in community development programs, which is largely absent. This is especially important when the goal is to influence or manipulate networks to achieve a desired result. This research can also be helpful for community development practice. Analysis of the communities' networks with organizations could serve as a useful diagnostic tool for community development professionals engaged in improving community capacity and network building. These networks can reveal social network capital of the communities, reflecting community capacity to meaningfully engage with others for local action.

This research observed network activity of communities in the program and comparison groups at one point in time ex-post the program intervention. It lacked baseline network activity for both groups of communities prior to program exposure, important for achieving greater methodological rigor. In addition, this research focused on network activity to assess the network building behavior of communities. The composition of the organizations contacted or the attributes of the contacted organizations are also equally meaningful. Also, the study is limited to the case of a single community infrastructure program, the context of which might be different from a community income generating program, for example, affecting network activity differently. Future research should address these limitations. Future research should also examine the questions discussed above that have important theoretical implications.

References

Agnitsch K., Flora J., Ryan V. (2006), "Bonding and bridging capital: The interactive effects on community action", *Community Development*, 37 (1), pp. 36-51.

Argyle M., Furnham A., Graham E.J.A. (1981), *Social situations*, Cambridge University Press, Cambridge.

Chaskin R.J., Brown P., Venkatesh S., Vidal A. (2001), *Building community capacity*, A. de Gruyter, New York.

Dale A., Newman L. (2008), "Social capital: A necessary and sufficient condition for sustainable community development?", *Community Development Journal*, 45 (1), pp. 5-21.

Granovetter M. (1985), "Economic action and social structure: The problem of embeddedness", *American Journal of Sociology*, 91(3), pp. 481-510.
Goldin I., Rogers H., Stern N. (2002), *The role and effectiveness of development assistance: Lessons from world bank experience*, The World Bank, Washington DC.
Lusher D., Koskinen J., Robins G. (*ed.*) (2013), *Exponential random graph models for social networks*, Cambridge University Press, New York.
Ostrom E. (2009), "A general framework for analyzing sustainability of social-ecological systems", *Science*, 325 (5939), pp. 419-422.
O'Toole L.J., Meier K.J. (2004), "Public management in intergovernmental networks: Matching structural networks with managerial networking", *Journal of Public Administration Research and Theory*, 14, pp. 469-494.
Pfeffer J., Salancik K.J. (1978), *The external control of organizations: A resource dependence perspective*, Stanford University Press, Stanford.
Provan K.G., Milward H.B. (1995), "A Preliminary theory of interorganizational network effectiveness: A comparative study of four mental health systems", *Administrative Science Quarterly*, 40, pp. 1-33.
Robins G., Pattison P., Kalish Y., Lusher D. (2007), "An introduction to exponential random graph (p*) models for social networks", *Social Networks*, 29, pp. 173-191.
Rossi P.H., Lipsey M.W., Henry G.T. (2019), *Evaluation – A systematic approach*, SAGE, Thousand Oaks, 8th ed.
Shrestha M.K. (2013), "Self-Organizing network capital and the success of collaborative public programs", *Journal of Public Administration Research and Theory*, 23 (2), pp. 307-329.
Shrestha M.K. (2019), "Network structure and spinoff effects in a collaborative public program", *Public Performance & Management Review*, 42 (4), pp. 976-1003.
Wang P. (2013), "Exponential random graph model extensions: Models for multiple networks and bipartite Networks", in D. Lusher, J. Koskinen, G. Robins (*eds.*), *Exponential random graph models for social networks: Theory, methods, and applications*, Cambridge University Press, Cambridge, pp. 115-129.
Wang P., Pattison P., Robins G. (2013), "Exponential random graph model specification for bipartite network – A dependence hierarchy", *Social Networks*, 35 (2), pp. 211-222.
Wang P., Robins G., Pattison P., Koskinen J. (2014), *MPNet: Program for the simulation and estimation of (p*) exponential random graph models for multilevel networks user manual*, Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne.
Wasserman S., Faust K. (1994), *Social network analysis – Methods and applications*, Cambridge University Press, New York.
Woolcock M. (1998), "Social capital and economic development: Toward a theoretical synthesis and policy framework", *Theory and Society*, 27 (2), pp. 151-208.
Woolcock M., Narayan D. (2000), "Social capital: Implications for development theory, research and policy", *The World Bank Research Observer*, 15 (2), pp. 225-249.

8. Italian top influencers on Twitter in COVID-19 time. A multiplex network analysis

by Vanessa Russo[*], Emiliano Del Gobbo[**], Lara Fontanella[***], Mara Maretti[****]

1. Introduction

COVID-19 is a primarily respiratory, infectious disease caused by a SARS-CoV-2 registered virus and belonging to the coronavirus family. Coronavirus mainly affects the lower respiratory tract and causes a series of flu-like symptoms, namely fever, cough, shortness of breath, muscle pain and tiredness; in some cases, it can be effectively asymptomatic. However, in severe cases, the virus can cause pneumonia, acute respiratory distress syndrome, sepsis, septic shock and even the death of the patient.

The SARS-Cov-2 virus is a new generation virus for which there is currently no vaccine or specific cure. All that can be done is the simple treatment of clinical symptoms.

The centre from which coronavirus spread was initially identified in the market in the city of Wuhan, the capital of Hubei province in China, as some patients developed pneumonia without a clear cause.

The first attributable report dates back to 31 December 2019, but the first patients with symptomatic disease had already appeared in early December.

On 21 February, two outbreaks of infections were reported in Italy, one in Lombardy and another in Veneto, with an initial number of 17 cases.

[*] Department of Legal and Social Sciences, University "G. d'Annunzio" of Chieti-Pescara, Chieti, Italy, russov1983@gmail.com.

[**] Department of Neuroscience and Imaging, University "G. d'Annunzio" of Chieti-Pescara, Chieti, Italy, emiliano.delgobbo@unich.it.

[***] Department of Legal and Social Sciences, University "G. d'Annunzio" of Chieti-Pescara, Chieti, Italy, lara.fontanella@unich.it.

[****] Department of Legal and Social Sciences, University "G. d'Annunzio" of Chieti-Pescara, Chieti, Italy, mara.maretti@unich.it.

Since then the epidemic has spread rapidly, and as of 17 June 2020, a total of 237,500 confirmed cases (including 34,405 fatal cases) had been recorded, most concentrated in the northern Italian regions, in particular in Lombardy.

The time span between 22 February 2020 and 8 June 2020 can be considered as one of the most crucial in Italian history. The critical issues not only concerned the spread of coronavirus and the management of the health emergency but above all the side effects of the pandemic: limitation of personal and religious freedoms, economic recession, crisis in international relations between EU countries and infodemic (Lazzerini & Putoto, 2020).

This contribution aims at analysing the dissemination of information in the Twittersphere during the COVID-19 health emergency in Italy, with a specific focus on the role of social actors in the social media of Twitter and how they are connected to each other.

Immediately after the declaration by the World Health Organization (WHO) of a COVID-19 health emergency, a real tsunami of information broke out, the main features of which can be summarised as: 1) high volume of information; 2) variety of sources of information (official and non-official); 3) multitude of information formats (infographics, videos, textual data, audio files); and 4) high viral potential.

In this regard, WHO Director-General Tedros Adhanom Ghebreyesus at the Munich Security Conference on 15 February 15 said, “We are not just fighting an epidemic; we are fighting an infodemic” (Zarocostas, 2020).

The term Infodemic has become a key word with respect to misleading information connected with the COVID-19 health emergency that has often become a contributing factor to collective social behaviours capable of altering the effectiveness of countermeasures put in place by governments.

These phenomena certainly include the spread of Fake News. However, the concept of infodemic, understood as “bad information”, is not exclusively attributable to false information or information from unofficial sources. Above all, it concerns the incorrect and not always consistent use of official media.

An example of this type concerns the news leak that was disseminated by CNN about the possible blocking of travel in the Lombardy region. The spread of that news, hours before the official communication of the Italian Prime Minister, caused a mass flight from the region to the southern regions increasing the risk of contagion (Cinelli *et al.*, 2020).

The WHO Director-General also stated that in such cases the development of infodemic phenomena is historically frequent. The feature that makes this phenomenon difficult to manage is the presence of Social Media (SM) (Zarocostas, 2020).

Social systems are huge information boosters, and information travels at a speed that is difficult to control. In addition, messages that travel within the SM have great visibility and are able to generate user collectives, who share and create affinity ties to become real opinion groups (van Dijck, 2013).

In this paper, we have decided to identify and analyse the characteristics of the communication and dissemination of information on COVID-19 on Twitter.

In particular, the analysis presented here is exploratory and aims to provide a first photograph of the characterisation of the flow of information related to the main production and dissemination agents of content on Twitter. The study is structured through the following research stages: a) analysis of the communication flow in relation to the main institutional communication events; b) identification of the main actors in the information networks on the basis of their productivity, their network of followers and the mentions they received and Retweets they obtained, which was made possible through the definition of a popularity index; c) selection of the 1,000 most popular users and, due to the complexity of the phenomenon, the analysis of the characteristics of the communication networks through a multilayer network analysis in which the layers of the analysis concerned different aspects, including the semantic affinities between the producers of information content; and d) a focus on the role of users (Activists, Bloggers, etc.) who conveyed disinformation content as well as conspiracy theorists.

The research aims at demonstrating how: a) the flow of communication on Twitter is sensitive and therefore closely connected to institutional communication; b) Activists and Bloggers who produce and disseminate official and unofficial information play an important role within the twittersphere; c) there are groups of users with specific aggregating characteristics that result from the sharing of specific social statuses, a political agenda or aggregating opinions that can be identified by specific languages and symbols (emoticons).

2. Theoretical background: the COVID-19 informational infoveillance on social networks

Although the pandemic emergence generated by COVID-19 is relatively recent, scholars in the Social Data Science area have promoted several studies on the communication flows in social media and social network environments. The observational infoveillance produced several useful and accessible databases (Chen *et al.*, 2020; Li *et al.*, 2020; Ziems *et al.*, 2020) and allowed various studies on the infodemic to be conducted (Cinelli *et al.*, 2020; Pennycook *et al.*, 2020; Gallotti *et al.*, 2020). The common goal of these papers is to analyse the characteristics of the viral online communication from different points of view: what are the differences between the patterns of junk and conspiracy contents; how much of the information is manipulated by human users and social bots; what are the space and temporal dynamics of online disinformation, etc.

In this latter aspect, the research by Singh *et al.* (2020) is interesting because it showed how an analysis of the flow of social communication on Twitter can be predictive of the epidemic's trend. The research also defined how disinformation and myths about COVID-19 are important but are not predominant subjects over others. Pulido *et al.* (2020) demonstrated how fake information is tweeted more (more produced) and that there is less retweeting of science-based evidence and fact-checking Tweets.

Other studies have focussed attention on how racial hate is spread during the pandemic and on the role of counter hate speech in the mitigation of the diffusion between social media users. Ziems *et al.* (2020) studied the evolution and spread of anti-Asian hate speech through Twittersphere and showed that 10.4% of users who produced hate content were bots. Bots and human users do not interact in isolated polarised communities but interact and influence other users. These hate environments influence other users, but research also shows that counter hate messages can discourage users from turning themselves into haters. The importance of bots on social media communication is also demonstrated by Ferrara (2020). In his study, he showed that automated social media accounts are active in the context of COVID-19 on Twitter and these bots are used for fuelling conspiratorial and ideological narratives in the United States. Specifically related to conspiracy theory, Ahmed *et al.* (2020) analysed the drivers in the United Kingdom of 5G COVID-19 conspiracy theory on Twitter. They identified two bigger network structures: an isolates group and a broadcast group as well as other smaller groups. The analysis also revealed that on a sample of 233 Tweets, 34.8% (n = 81) contained conspiracy theory on the relation of 5G and COVID-19, 32.2% (n = 75) contained counter-conspiracy Tweets and 33.0% (n = 77) contained neutral Tweets. This suggests that, although the topic attracted high volume, only a few users genuinely believed the conspiracy. More specifically, Mourad *et al.* (2020) analysed the reliability of the content conveyed on Twitter. The research showed the severe impact of information that misleads people and the spread of unreliable information on Twittersphere during the COVID-19 time as well as the importance of studying the characteristics of the opinion leader on social media.

All these studies are very interesting and shed light on the info-pandemic phenomenon in COVID-19 time. However, with the aim of understanding the Italian case study, it is useful to start from the analysis of Twitter influencers to detect the main communities and their characteristics and to focus attention on the identification of the main content producers and the main content shearing users.

3. Data and method

The choice to use Twitter is due both to the streamlined and dynamic structure of social media and to the possibility of easy access to a significant amount of public content. Twitter presents itself as a relationally asymmetric microblogging platform. The diffusion of content, called Tweets, must be summarised in 280 characters, and the links between the user profiles are not necessarily two-way but can be non-reciprocal and grouped by Friends and Followers (Bentivegna, 2014, 2015; Murthy, 2018; van Dijck, 2011). These basic features make Twitter a much slimmer and more flexible SM than Facebook; in fact, Hagan defines it as "a sort of adrenalized Facebook" (van Dijck, 2013, p. 70).

Twitter offers each user a generally public and very concise personal space in the personal description, while the flow of communication is divided into three fundamental actions: 1) the production of content (Tweet); 2) the sharing of content with other users (Retweet); and 3) the production of content by connecting users to specific concepts (Mention). Beginning in 2007, a content cataloguing system that was already in use in the context of Internet Relay Chat (IRC) was introduced, marked with the # symbol and a defined hashtag with the popularity conveyed by the Trending Topic algorithm. The hashtag is a fundamental part of the syntax of Twitter and has a dual function, one manifest and the other latent. The first is attributable to the function of the hashtag itself, that is, to make the Tweets related to a certain topic identifiable; the second is latent and aimed at outlining the audiences that form ad hoc around specific topics (Bentivegna, 2015).

3.1. Collecting the data

The data collection phase from Twitter was carried out by establishing a streaming connection to the social media API. Through this access system, it was possible to acquire Tweets in real time over a certain time period.

The data collection process was organised into three phases.

The first phase began on 22 February 2020, starting with the collection of the most popular hashtags on the COVID-19 theme (#coronavirus, #covid-19, #coronavirusitalia, #covid19italia). After the start of the lockdown across the country (dpcm of 8 March 2020), #iorestoacasa was added. In the third and final phase corresponding to the beginning of phase 2 (dpcm 26 April 2020) a further hashtag #Fase2 (#phase2) was added.

The Tweets dataset was collected using Socialgrabber, an online service platform, that provides a user-friendly GUI to use the publicly available Twitter Streaming API (Access Programming Interface).

The platform allows the collection of data scheduling and data acquisition jobs and the setting of filtering parameters, and it allows them to be exported in several formats. We exported data in the "JSON Lines" format (http://jsonlines.org/). Twitter provides several fields of data for each Tweet which refers to the user, the text, the time and the entities occurring in each Tweet[1] (hashtags, mentions, URLs). Twitter Streaming API provides Retweets together with the Tweets. A Retweet is when a user just republishes a post of another user. We processed the collected dataset using several Python scripts, and we developed a script to collect all the friends of the top 1,000 users by using the Twitter API user, ranked according to a system that will be better explained in Section *Methods & Data processing*. It is worth noting that the term friend refers to the users that a given user follows on Twitter.

3.2. Composition of the dataset

A total of 7,306,469 Tweets and Retweets were collected through the data acquisition phase. Compared to the complete data set, Tweets were 2,021,313 while Retweets were 5,285,156.

The flow of Tweets is signalled by frequency peaks (Fig. 2), which are particularly high in some periods characterised by significant institutional decisions. The frequency curve shown in Figure 2 shows an increase in data in correspondence with: the first Decree-Law "Urgent measures for the containment and management of the COVID-19 epidemiological emergency", of the first Decree-Law (8 March and 11 March 2020) and the "Cura Italia" Decree (Decree-Law of 17 March 2020, n. 18: "Strengthening measures of the National Health Service and economic support for families, workers and companies related to the epidemiological emergency from COVID-19").

[1] For an overview of tweet data fields, visit https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/tweet-object.

Fig. 1 – Temporal distribution of Tweets and Retweets

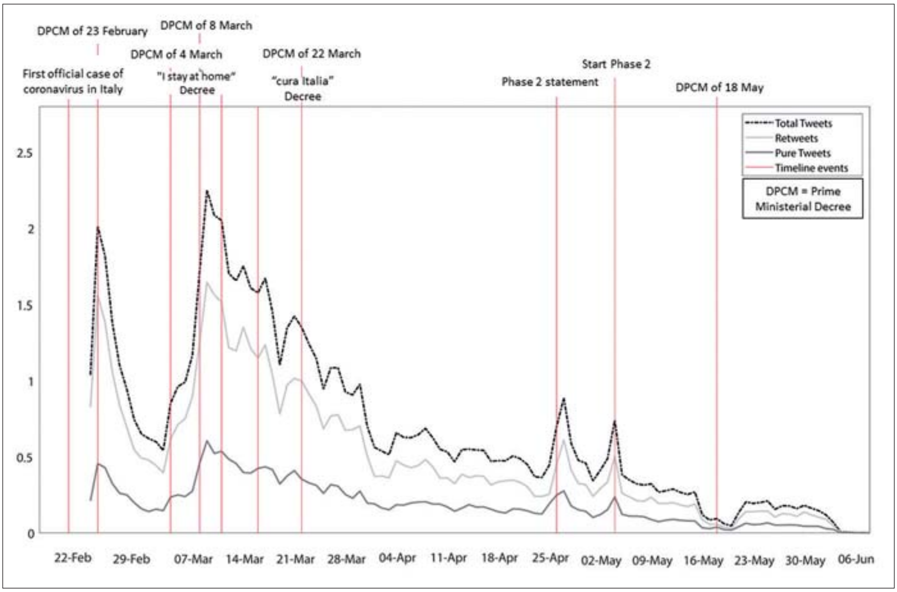
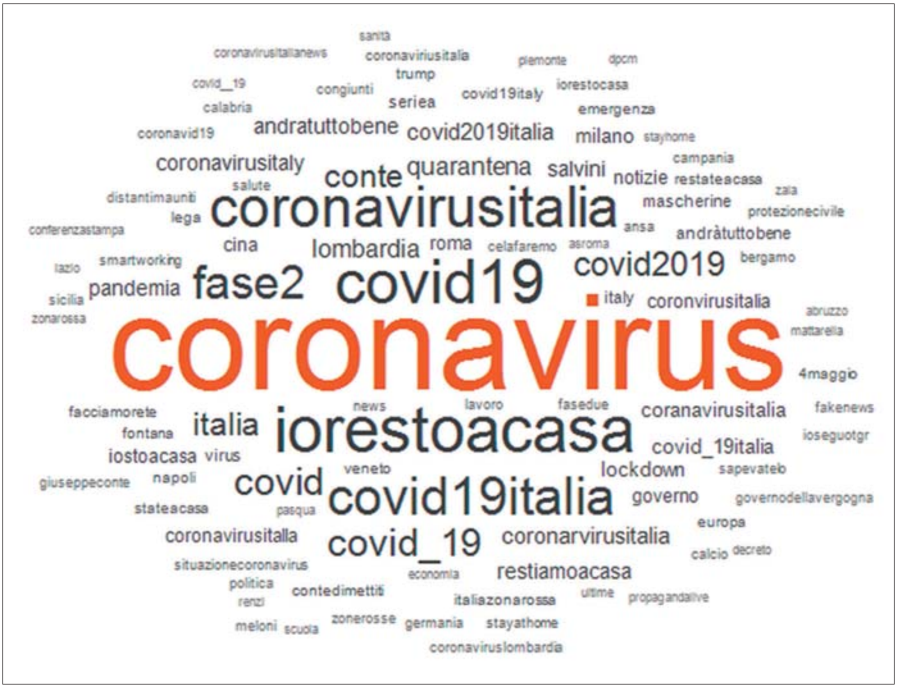


Fig. 2 – Hashtags' WordCloud



From the content point of view, a hashtags list which was needed throughout the dataset emerged (Fig. 2).

Specifically, it is possible to identify (Fig. 2) a list of frequent terms which are particularly descriptive of the situation and of the social actors involved.

The analysis of the occurrences shows:

– a subset of hashtags used as synonyms for coronavirus (#coronavirus, #covid19, #coronavirusitalia, #covid19italia, #covid_19, #covid, #covid2019, #covid2019italia, #coronavirusitaly);
– a semantic area of the keywords that describes the lockdown (#iorestoacasa, #phase2, #pandemic, #quarantine, #andràtuttobene, #lockdown, #iostoacasa, #notizie, #mascherine); and
– a group of words representing some of the social actors and institutions that characterise the whole story: #conte, #salvini, #italia, #lombardia, #milano, #governo, #roma, #cina and #seriea.

3.3. Methods & Data processing

There is a vast scientific literature on the applications of Social Network Analysis in the digital environment of Twitter (Ediger *et al.*, 2010; Himelboim *et al.*, 2017; Chatfield & Brajawidagda, 2012; Khonsari *et al.*, 2010; Yoon & Park, 2014; Tremayne, 2014).

The social media structure easily lends itself to SNA because it presents a network of three-dimensional relationships: 1) asymmetry between users (friends and followers); 2) asymmetry between users and concepts (@mention); and 3) symmetry between concepts and concepts (#hashtag). These three dimensions are related to the entire Twitter sphere.

The first two can be considered as social networks because they connect agents, while the third is a semantic network because it allows you to map the links of co-occurrence between the different hashtags.

In addition, Twitter content is very short (each Tweet has a maximum of 280 characters) and can be easily analysed using Semantic Social Network Analysis, a collection of search techniques that treat a word like a node in a network and a semantic relation (in terms of concordance) between words as a social relationship that connects those words (Chung & Park, 2010; Jung & Park, 2015).

In order to exploit all the potential offered by the Twitter structure, it was decided to analyse the behaviour of twittersphere agents by analysing in parallel the social and semantic dimensions. In particular, four networks have been built between users of the twittersphere, two that exploit social ties based, respectively, on Friends/Followers and Mentions, and two that take into consideration the semantic ties determined by the sharing of a shared word and hashtag. These networks can be analysed by exploiting the potential of multilayer networks.

Given the size of the dataset, the field of analysis was restricted to a subset of 1,000 users chosen in relation to their potential to be influential in the Twitter discussion sphere.

For this reason, the data processing phase was organised according to two main objectives: 1) the extraction of information to detect Opinion Leader users throughout the dataset; 2) the construction of the adjacency matrices to map their semantic and social relationships.

Through a script written with the Python programming language, a data frame composed of all the users who produced the content was extracted. Information was also extracted for: 1) the number of Followers for each user; 2) the number of posts produced with the labels identified; 3) the content Retweet value; and 4) the mention number. All the information collected was used to build an index capable of detecting a list of a thousand particularly influential users, classified in relation to their social role in the twittersphere.

The dataset was then further processed in order to extract four adjacency matrices using the opinion leaders as nodes and, respectively, friends, mentions, word sharing and hashtag sharing.

3.4. The Opinion Leaders

Several studies have been conducted with the aim of identifying influential user profiles on Twitter. Usually this research explores influence metrics to detect influential users as opinion leaders, innovators (Chai *et al.*, 2013), authoritative actors (Bouguessa & Romdhane, 2015) or prestigious individuals (Gayo-Avello, 2013). Some methods that aim at evaluating the influence of Twitter users are based upon individual profile parameters or interactions with other online users. Other influence measurement techniques integrate graph theory with structural parameter (Nagamoti *et al.*, 2010; Noro *et al.*, 2013; Pal & Counts, 2011; Bigonha *et al.*, 2012; Cappelletti & Sastry, 2012; Bigonha *et al.*, 2012; Kundu *et al.*, 2011; Jain & Sinha, 2020; Choi & Park, 2015; Lorentzen & Nolin, 2017).

In order to circumscribe the field of analysis and select influential users, a synthetic index was developed that is able to bring out the relevant social actors in the reference sphere of discussion.

The choice of indicators to be examined is based in particular on the methodological proposals of Cha *et al.* (2010), Choi (2015), Marchetti in Bentivegna (2014) and Leavitt *et al.* (2009).

Within Twittersphere, the number of followers of a user is absolutely reductive and the risk exists of incurring the so-called Million Follower Fallacy as a reference point for defining an opinion leader (Cha *et al.*, 2010).

In fact, the number of Followers cannot be considered as a value capable of exhaustively indicating the figure of the influencer because there are agents who have a certain number of Followers exclusively by virtue of their social role. Therefore, they do not really have an active audience. Several empirical studies have elaborated synthetic indexes which consider follower, retweets and mentions (Cha *et al.*, 2009; Choi, 2015; Marchetti in Bentivegna, 2014; Leavitt *et al.*, 2009).

Cha, in particular, defines as effective: in-degree influence, retweet influence, mention influence. In particular, the in-degree influence measures the size of the audience of each user; retweet influence instead measures the ability of each user to generate viral content; and the mention influence implies the ability of each user to be involved in other conversations (Cha *et al.*, 2010).

Starting from Cha's work (Cha *et al.*, 2010), an index has been structured here that takes into consideration for each user: the number of posts produced; the maximum value of Retweets; the sum of the Retweets; the number of followers; and the number of mentions.

The indicators were subsequently normalised by heating them between 0 and 1 through the min-max standardisation process: Z = ((x-min (x)) / (max (x) -min (x). The indicators were aggregated through the arithmetic mean.

3.5. Multiplex Network Analysis

As detailed in Section 3.3, we take into account different links among the identified influential agents or opinion leaders; therefore, it is natural to model this complex interaction system as a multiplex network (Kivelä *et al.*, 2014) which is a combination of individual networks coupled through links that connect each node in one network to itself in other networks.

The adjacency matrices of our analysis were derived differently for the different links. In particular, for the friend/follower and the mention networks, we considered direct unweighted graphs. On the other hand, we built directed weighted graphs for the networks based on the Tweet contents and on the hashtags. For these latter networks, the weights measured the lexical similarity or the similarity in the word or hashtag usage. In more detail, the weight of the link between any two opinion leaders was obtained by rescaling the number of words (or hashtags) in common (i.e., the similar content) by the product of the number of words (or hashtags) used by each of them (i.e., user lexical diversity). In this way, a weight equal to 0 indicates that there are no words (hashtags) in common, while a weight equal to 1 indicates that the two opinion leaders use the same vocabulary (or the same set of hashtags).

The first aim of our analysis is to decompose the networks into subunits or communities consisting of highly interconnected nodes. To perform community detection on the multiplex structure, we adopted the generalised Louvain method (Mucha *et al.*, 2010) which generalises the determination of community structure via quality functions to multi-slice networks that are defined by coupling multiple adjacency matrices.

For each slice s, the network structure is represented by adjacencies A_{ijs} between nodes i and j; in addition the single slice adjacency matrices are connected through interslice couplings C_{jsr} that connect node j in slice r to itself in slice s. The strength of each node individually is given by $k_{is} = \Sigma_j A_{ijs}$ in each slice and by $g_{is} = \Sigma_r C_{isr}$ across slices; therefore the multi-slice strength is $k_{is} = K_{is} + g_{is}$.

The multi-slice generalisation of modularity is given by

$$Q_{multislice} = \frac{1}{2\mu}\sum_{ijsr}\left[\left(A_{ijs} - \gamma_s \frac{k_{is}k_{js}}{2m_s}\delta_{sr}\right) + \delta_{ij} C_{jsr}\right]\delta(g_{is}, g_{jr})$$

where $m_s = \Sigma_i k_{is}$ is the total strength for slice s, $2\mu = \Sigma_{is} k_{is}$ and δ is the Kronecker delta.

This generalised formulation of the Louvain method incorporates resolution parameters which influence the number of identified communities: γs are slice specific resolution parameters, while the interslice resolution parameter is incorporated in the elements of C_{js}r which are assumed to take binary values {0, ω}.

For our analysis, we specified the slice specific resolution parameters equal to 1, as in the classical Lovain approach, and the interslice coupling equal to 0.5. The analysis was performed in Matlab using the GenLouvain toolbox (Jeub *et al.*, 2011, 2019).

Another aim of the analysis is related to the identification of the importance of entities in the network based on the social links. In directed networks, nodes can have very different roles if we consider only the input or output links. The idea of the hyperlink-induced topic search (HITS) approach, also known as hubs and authorities' algorithm (Kleinberg, 1999), is to assign two scores to each node: a hub centrality, which takes into account the role of the node in sending links, and an authority centrality, measuring the capacity of the node to receive links. Following the same approach of eigenvector centrality, the importance as authority depends on the relevance of the hubs that send the incoming links and, vice versa, important hubs give more weight as authorities to the receiver nodes.

To assign hub and authority scores, we adopted the approach proposed by Taylor and colleagues (2019) in the multiplex context. This approach involves coupling centrality matrices that are associated with individual layers into a larger supra-centrality matrix and studying its dominant eigenvector to obtain joint, marginal and conditional centralities. The computation of the supra-centrality matrix is based on the same coupling process described for the community detection phase. In particular, defining with $\{B_{(s)}\}$, $s = 1,\ldots,T$ the set of centrality matrices for the slices of a multiplex and with C the interlayer adjacency matrix, the family of supra-centrality matrices $\tilde{S}(\omega)$, parameterised by the interlayer-coupling strength ω, can be built as $\tilde{S}(\omega) = \tilde{B} + \omega\tilde{C}$, where $\tilde{B} = diag((B^{(1)}, \ldots, B^{(T)})$ and $\tilde{C} = C \otimes I$, with I the identity matrix and ⊗ the Kronecker product.

Given a supra-centrality matrix $\tilde{S}$, the dominant eigenvector (i.e., the eigenvector associated to the largest eigenvalue) of $\tilde{S}'\tilde{S}$ provides the hub scores, while the dominant eigenvector of $\tilde{S}\tilde{S}'$ gives the authority scores. From this spectral decomposition, we obtained the joint centrality measures for each node in each layer. The sum of the centrality measures across layers gave the marginal centrality.

The implementation of the computation was performed in MATLAB adapting the code developed by Taylor and available at https://github.com/taylordr/Supracentrality.

4. Results

The research results are ordered according to a prerequisite. The identification and classification of opinion leaders has made it possible to highlight all the nodes of the graph. Through the mapping phase of the multilayers, the different communities of nodes have been identified, interpreted in relation to the quality of the nodes that make them up.

Finally, we played the role of opinion leaders by extracting the values of Hub Centrality and Authority Centrality from multiplexes.

4.1. Opinion Leaders

A subset of 1,000 influential users was extracted through the popularity index, and the identified nodes were classified according to their role within the social media Twitter.

For the classification, we developed the following labels:

– Politician: user pages managed by Politicians;
– Political Party/Movement-pages of political parties and movements;
– Institutional: official profiles of institutional bodies;
– News: profile pages of Newspapers, Press Agencies, Newscasts, Web News;
– Media: user profiles of television, radio and web channels.
– Magazine: user profiles of entertainment and in-depth information magazines;
– Scientist: profile pages of experts in a specific scientific field;
– Activist–profile of users who define themselves as activists or who have a specific reference agenda;
– Journalist: user profile of information professionals;
– Satire: pages of satire;
– Blog/Blogger: profiles directly linked to a blog or blogger;
– NEP–Pages no longer existing;
– Association–user profiles of associations;
– Public Figure: profiles of public figures and celebrities;
– Users: pages of active users on Twitter who spread generic content;
– Sports: user pages of sports clubs; and
– Others: the other item includes all the pages not classified in the previous categories.

Compared to labels, we have detected through the classification process a very heterogeneous universe of users (Table 1).

The largest group (163 users) is that of users classified as Activists. The group differs from the Users because the first is composed of Twitter users active on various themes and with an often-unspecified ideological orientation, while the activists are users who declare a specific membership and produce content with respect to a specific agenda of references. The Activist group is made up of three subcategories: populist right, progressive left, LGBT activists.

Users with a right orientation are easily identifiable: in the description of the username there is the Italian flag accompanied by a green or black heart. The former defines an ideological closeness to the League, the latter to Brothers of Italy. The profiles of right-wing Activists are often accompanied by anti-immigration and anti-European hashtags.

Tab. 1 – Opinion leaders

Classification	*User*	*Tweets*
Activist	16.3%	15.6%
Journalist	12.6%	5.7%
News	11.0%	34.5%
Politician	9.8%	1.6%
User	9.1%	5.6%
Institution	8.9%	4.8%
Media	6.2%	3.8%
Blog/Blogger	6.1%	11.4%
Public Figure	4.8%	0.5%
Sport	3.0%	4.6%
NEP	2.8%	5.7%
Association	2.5%	2.0%
Magazine	2.0%	2.0%
Political Party/Movement	1.6%	0.4%
Scientist	1.2%	0.6%
Satire	1.1%	1.0%
Other	1.0%	0.2%
Total	100.0%	100.0%

The left-wing user profiles have the Italian flag flanked by the European flag and are accompanied by descriptions and hashtags specifically anti-fascist and pro-European. Some of the leftist activists accompany their name with the #facemorete hashtag. It represents a Twitter network made up exclusively of leftist activists.

Finally, LGBT activists have a user profile accompanied by the rainbow emoticon and they define themselves in the user description as LGBT activists.

Then there are "Journalists": users who work for press bodies and news or profile pages of press agencies, newspapers and online portals. The Journalists and News nodes differ from Blog/Blogger (61 users) because the latter refers to unofficial information portals and their authors.

Finally, in the information field we have classified some profiles as magazines (20 users). These are the user pages of rotogravure and in-depth information sources.

The third largest group is that of politicians or political figures who hold institutional positions or who use Twitter as a space for communication and political propaganda. Together with the group of politicians, there is the "Political party/Movement" group, which includes official profiles of political parties and movements.

The smallest groups are: Sport, Satire, Scientist, PN and Others. Sport includes the user profiles of sports clubs, while Satire the portals of Satire. With the Scientist label, we have collected users with specific expertise in the medical-scientific field. Finally, the profiles that no longer exist or have been closed have been labelled with PN and the profiles that do not fall into any of the previous classifications with Others.

4.2. Network and Community detection

The representation of the Multiplex – using Gephi software (Bastian *et al.*, 2009) – was done through 4 graphs, two semantic and two social (Fig. 3, Fig. 4). To identify the nodes of the graph, the classification of Opinion Leaders identified in the previous paragraph was used, and the nodes were coloured according to the labels. The nodes of the network were weighted in relation to the popularity index shown in the Methods and Data processing paragraph. Finally, the clusters obtained through multiplex community detection were represented in the various networks.

The friendship ties graph (Friends/Followers) is composed of 1,000 nodes with 90,017 connections and 29 communities, 5 of which are relevant and dense in content (Fig. 3).

Cluster n. 3 (204 nodes) has many activists and politicians within it, in the clique there are several users with political ideas of the right populist, including the two political representatives Matteo Salvini and Giorgia Meloni and the newspaper Libero. The cluster also includes the page of the Radio Savana blog, often identified as a source of fake news and junk information.

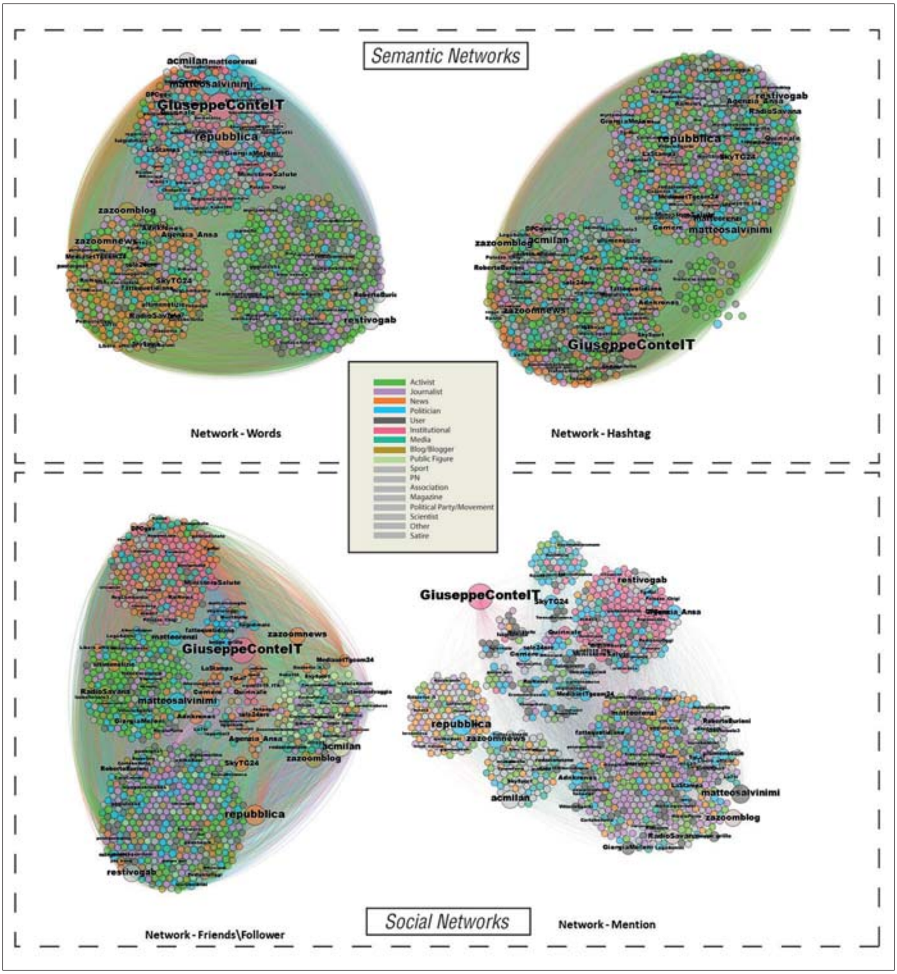
The community n. 5 is the most numerous (377 knots) and is mainly composed of Journalists, Activists and Politicians. Differently from n. 3, clique 5 has a left-wing orientation in fact in this cluster there are many members of the Democratic Party including Laura Boldrini as well as many moderate and radical left-wing activists.

Cluster 9 (198 knots) is characterised by News, Institutional pages and Associations. In this cluster, the pages of the various ministries involved in the health emergency (Health, Economy, Defence, Education) and some Italian regions were found.

The community n. 4 (138 knots) contains pages of users from the world of entertainment and publishing plus sports and sports associations. In this cluster, there are also the profile pages of Barbara D'Urso and of Journalists and Bloggers who mainly deal with entertainment information.

Finally, in cluster 1 (29 nodes), there are many pages of Politicians and Political parties. The political figures in the community are all members of the 5-star movement; in fact the pages of Giuseppe Conte, Luigi Di Maio, Virginia Raggi, Beppe Grillo are included. In the group, there are also the pages of the *Fatto Quotidiano* newspapers and journalist Marco Travaglio.

Fig. 3-4 – Social and Semantic Networks



The Mention network, on the other hand, is made up of 11,524 links and 34 communities, 5 of which are dense with knots (Fig. 3). The largest group is the n. 3 composed largely of Activists, Journalists and Politicians of different orientations (447 knots). In this cluster, in fact, there are nodes belonging to the pages of Matteo Salvini and Giorgia Meloni, Matteo Renzi and Beatrice Lorenzin.

Communities 19, 1 and 4 are analogous to the Friends/Followers (FF) network. Group 19 has a very similar composition to clique n. 9 of the previous network; in fact, both have a purely institutional matrix and are composed of the ministries of health, education, economic development and are from the same official pages of the Italian Regions. Cluster 1, on the other hand, has several elements in common with group 1 of the FF network, both of which have the official profile of Giuseppe Conte and Luigi Di Maio. The community n. 4 is characterised by nodes belonging to the world of entertainment and sport in analogy to group 4 of the FF network.

Of particular relevance is group 14 composed of many nodes belonging to the world of information. In fact, the most frequent nodes are labelled as News, Journalist and Magazine.

The Words network (Fig. 4) is made up of 489,125 connections and 19 communities, 3 of which are larger (Fig. 4). Group 5 is the most numerous (409 knots) and includes nodes classified as Activists, Journalists, Users and Public Characters.

The community n. 19 presents 314 users with the prevalence of these categories: Politicians, Political parties, Institutions, Media and Associations. In cluster n. 3 there are the News and Sport nodes.

In analogy with the Words network, the Hashtag network also consists of 19 communities, n. 19 (528 nodes) have a prevalence of Politician nodes. Group 3 (424 knots) has a prevalence of knots labelled as Journalist, Institutional, Media, Public Figures and Sport.

Within the Hashtag network, community n. 5 (48 Nodes) is interesting, mainly composed of Activist and Blogger nodes. In the clique, there are many users with a specific political orientation, among them are some nodes with a higher popularity index and contents that are admittedly close to the populist right.

4.3. Hub & Authorities

By considering only the two directed graphs based on the social links, friends/followers and mentions, we derived the multiplex Hub and Authority scores (Taylor *et al.*, 2019) of the identified influential agents. Hubs are nodes which point to many nodes of the type considered important. Authorities are these important nodes. From this comes a circular definition: good hubs are those which point to many good authorities and good authorities are those pointed to by many good hubs (Kleinberg, 1999). Table 2 shows the composition of the first 50 nodes and of the nodes ranked from 50 to 100 in terms of hub score and authority scores in terms of the qualitative classification described in Section 4.1.

As Table 2 shows, the first 50 nodes with a higher level of Hub Centrality belong to the Journalist, Media and Politician categories.

Tab. 2 – Distribution of Hubs & Authorities

	Hubs		*Authorities*	
	1-50	*50-100*	*1-50*	*50-100*
Activist	8.00%	18.00%	–	4.00%
Association	2.00%	–	–	–
Blog/Blogger	8.00%	8.00%	2.00%	4.00%
Institutional	8.00%	2.00%	12.00%	6.00%
Journalist	28.00%	36.00%	22.00%	28.00%
Magazine	–	–	–	–
Media	12.00%	2.00%	6.00%	10.00%
News	6.00%	8.00%	34.00%	18.00%
Other	2.00%	–	–	–
PN	2.00%	–	–	–
Political Party/Movement	2.00%	–	4.00%	4.00%
Politician	14.00%	22.00%	20.00%	24.00%
Public Figure	4.00%	–	–	–
satire	2.00%	–	–	–
Scientist	–	–	–	2.00%
Sport	–	–	–	–
User	2.00%	4.00%	–	–
Total	100.00%	100.00%	100.00%	100.00%

The first 50 nodes with a higher level of Authority Centrality are News, Journalist, Politician, Political Party/Movement and Institutional.

The analysis carried out shows that Journalist and Politician users are simultaneously both elements with a high degree of Hub Centrality and Authority Centrality, while Blogger and Blog and Satire pages have a privileged role specifically in the dissemination of information.

The agents with a high value exclusively of Authority Centrality are the information pages (News) and the user profiles of Political parties and movements.

5. Conclusions

The exploratory case study of the twittersphere in COVID's time allows us to draw some preliminary conclusions. First of all, taking into consideration the most influential communication agents (the first 1,000), it is clear that there is a varied environment of polarised users, especially around press agencies and institutional users. In fact, they are the main producers of information (Authorities in Table 2). This information is conveyed and commented (partly politically polarised) by Bloggers, Journalists, Politicians, Media and Satire (Hub in Table 2). The link between information flows and Institutional and Journalistic communication is evident in the trend shown in Figure 1. The top influencers are, therefore, institutions, politicians and, in particular, press agencies. The main spreaders are Bloggers, Media, Journalists, and satirical commentators. These groups appear to have a political polarisation function. In fact, another interesting aspect emerging from the analysis is represented by the existence of groups of users with specific aggregating characteristics from a political point of view. These communities are characterised by the sharing of a political agenda polarised on the right (cluster no. 3 Fig. 3) or on the left (cluster no. 5 Fig. 3). They can be identified by specific languages and symbols (emoticons) that distinguish them politically.

These results seem to confirm what emerges from other research on different national contexts that have dealt with infodemics in the twittersphere. Some of them point out that, although the twittersphere is populated by disinformation and myths about COVID-19, they are not predominant compared to those with official information content from accredited press agencies and institutions (Singh *et al.*, 2020). The same way results are presented in Pulido *et al.* (2020). In fact, even in our case study, the most retweeted contents have reliable sources.

With respect to future studies, it will be necessary to deepen the virulence of counter-information contents related to emerging conspiracy theories, as well as to analyse the role of social bots in communicating the health emergency in the twittersphere in Italy.

References

Ahmed W., Vidal-Alaball J., Downing J. and Seguí, F.L. (2020), "COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data", *Journal of Medical Internet Research*, 22, 5, e19458.

Bastian M., Heymann S., Jacomy M. (2009), "Gephi: An open source software for exploring and manipulating networks", *Third International AAAI Conference on*

Weblogs and Social Media, Menlo Park, CA, AAAI Publications, https://aaai.org/ocs/index.php/ICWSM/09/paper/view/154/100.
Bentivegna S. (2015), *Colpi di tweet. La politica in prima persona*, il Mulino, Bologna.
Bentivegna S. (a cura di) (2014), *La politica in 140 caratteri: Twitter e spazio pubblico*, FrancoAngeli, Milano.
Bigonha C., Cardoso T.N., Moro M.M., Gonçalves M.A., Almeida V.A. (2012), "Sentiment-based influence detection on Twitter", *Journal of the Brazilian Computer Society*, 18, 3, pp. 169-183.
Bouguessa M., Romdhane L.B. (2015), "Identifying authorities in online communities", *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6, 3, pp. 1-23.
Cappelletti R., Sastry N. (2012, December), "Iarank: Ranking users on twitter in near real-time, based on their information amplification potential", in *International Conference on Social Informatics*, IEEE, Alexandria (VA), pp. 70-77.
Cha M., Haddadi H., Benevenuto F., Gummadi K.P. (2010, May), "Measuring user influence in twitter: The million follower fallacy", in *Fourth International AAAI Conference on Weblogs and Social Media*, AAAI Press, Menlo Park (CA).
Chai W., Xu W., Zuo, M., Wen X. (2013, June), "ACQR: A Novel Framework to Identify and Predict Influential Users in Micro-Blogging", in *Proceedings of the Pacific Asia Conference on Information Systems (PACIS)*, Association for Computing Machinery, London, p. 20.
Chatfield A., Brajawidagda U. (2012), "Twitter tsunami early warning network: a social network analysis of Twitter information flows", in J.W. Lamp (*eds.*), *23rd Australasian Conference on Information Systems*, ACIS, Geelong (AUS), pp. 1-10.
Chen E., Lerman K., Ferrara E. (2020, March), "#COVID-19: The first public coronavirus Twitter dataset", *arXiv Prepr*, arXiv2003.07372, text available at: http://arxiv.org/abs/2003.07372.
Choi S. (2015), "The two-step flow of communication in Twitter-based public forums", *Social Science Computer Review*, 33, 6, pp. 696-711.
Choi S., Park H.W. (2015), "Networking interest and networked structure: A quantitative analysis of Twitter data", *Social Science Computer Review*, 33, 2, pp. 145-162.
Cinelli M., Quattrociocchi W., Galeazzi A., Valensise C.M., BrugnoliE., Schmidt, A.L., Scala A. (2020), "The COVID-19 social media infodemic", *arXiv preprint arXiv:2003.0500*.
Ediger D., Jiang K., Riedy J., Bader D.A., Corley C., Farber R., Reynolds W.N. (2010, September), "Massive social network analysis: Mining twitter for social good", in *39th International Conference on Parallel Processing*, IEEE, San Diego (CA), pp. 583-593.
Gallotti R., Valle F., Castaldo N., Sacco P., De Domenico M. (2020, April), *Assessing the risks of "infodemics" in response to COVID-19 epidemics*, available: http://arxiv.org/abs/2004.03997.
Gayo-Avello D. (2013), "Nepotistic relationships in Twitter and their impact on rank prestige algorithms", *Information Processing & Management*, 49 (6), pp. 1250-1280.

Himelboim I., Smith M.A., Rainie L., Shneiderman B., Espina C. (2017), "Classifying Twitter topic-networks using social network analysis", *Social media+ society*, 3 (1), 2056305117691545.
Jain S., Sinha A. (2020), "Identification of Influential users on Twitter: A Novel Weighted Correlated Influence Measure for COVID-19", *Chaos, Solitons & Fractals*, 110037.
Jalali S.M.J., Park H.W.@ (2017), "Conversations about open data on Twitter", *International Journal of Contents*, 13 (1), pp. 31-37.
Jeub L.G.S., Bazzi M., Jutla I.S., Mucha P.J (2011-2019), "A generalized Louvain method for community detection implemented in MATLAB", http://netwiki.amath.unc.edu/GenLouvain, https://github.com/GenLouvain/GenLouvain.
Jung K., Park H.W.@ (2015), "A Semantic (TRIZ) Network Analysis of South Korea's 'Open Public Data' Policy", *Government Information Quarterly*, 32 (3), pp. 353-358.
Khonsari K.K., Nayeri Z.A., Fathalian A., Fathalian L. (2010, August), "Social network analysis of Iran's Green movement opposition groups using Twitter", in *International conference on advances in social networks analysis and mining*, IEEE, Odense (DK), pp. 414-415.
Kivelä M., Arenas A., Barthelemy M., Gleeson J.P., Moreno Y., Porter M.A. (2014, September), "Multilayer networks", *Journal of Complex Networks*, 2 (3), pp. 203-271.
Kleinberg J.M. (1999), "Authoritative sources in a hyperlinked environment", *Journal of the ACM* (*JACM*), 46 (5), pp. 604-632.
Kundu S., Murthy C.A., Pal S.K. (2011, June), "A new centrality measure for influence maximization in social networks", in *International Conference on Pattern Recognition and Machine Intelligence*, Springer, Berlin-Heidelberg, pp. 242-247.
Lazzerini M., Putoto G. (2020), "COVID-19 in Italy: momentous decisions and many uncertainties", *The Lancet Global Health*, 8 (5), e641-e642.
Leavitt A., Burchard E., Fisher D., Gilbert S. (2009), "The influentials: New approaches for analyzing influence on twitter", *Web Ecology Project*, 4 (2), pp. 1-18.
Li J., Xu Q., Cuomo R., Purushothaman V., Mackey T. (2020, April), "Data Mining and Content Analysis of Chinese Social Media Platform Weibo During Early COVID-19 Outbreak: A Retrospective Observational Infoveillance Study", *JMIR Public Heal. Surveill.*, 6 (2), p.e 18700.
Lorentzen D.G., Nolin J. (2017), "Approaching completeness: Capturing a hashtagged Twitter conversation and its follow-on conversation", *Social Science Computer Review*, 35 (2), pp. 277-286.
Mourad A., Srour A., Harmanani H., Jenainatiy C., Arafeh M. (2020), "Critical Impact of Social Networks Infodemic on Defeating Coronavirus COVID-19 Pandemic: Twitter-Based Study and Research Directions", *arXiv preprint arXiv:2005.08820*.
Mucha P.J., Richardson T., Macon K., Porter M.A., Jukka-Pekka O. (2010), "Community Structure in Time-Dependent, Multiscale, and Multiplex Networks", *Science*, 328, pp. 876-878.

Murthy D. (2018), *Twitter*, Polity Press, Cambridge.
Nagmoti R., Teredesai A., De Cock M. (2010, August), "Ranking approaches for microblog search", in *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, IEEE, Toronto (ON), Vol. 1, pp. 153-157.
Noro T., Ru F., Xiao F., Tokuda T. (2013), "Twitter user rank using keyword search", *Information Modelling and Knowledge Bases XXIV. Frontiers in Artificial Intelligence and Applications*, 251, pp. 31-48.
Pal A., Counts S. (2011, February), "Identifying topical authorities in microblogs", in *Proceedings of the fourth ACM international conference on Web search and data mining*, Association for Computing Machinery, New York, pp. 45-54.
Pennycook G., McPhetres J., Zhang Y., Rand D. (2020), "Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention", *PsyArXiv*, working paper, pp. 1-24.
Pulido C.M., Villarejo-Carballido B., Redondo-Sama G., Gómez A. (2020), "COVID-19 infodemic: More retweets for science-based information on coronavirus than for false information", *International Sociology*, 0268580920914755.
Pulido C.M., Villarejo-Carballido B., Redondo-Sama G., Gómez A. (2020), "COVID-19 infodemic: More retweets for science-based information on coronavirus than for false information", *International Sociology*, 0268580920914755.
Rovetta A., Bhagavathula A.S. (2020), "COVID-19-related web search behaviors and infodemic attitudes in italy: Infodemiological study", *JMIR Public Health and Surveillance*, 6 (2), e19374.
Singh L., Bansal S., Bode L., Budak C., Chi G., Kawintiranon K., Paddena C., Vanarsdalla R., Vragad E., Wang Y. (2020), "A first look at COVID-19 information and misinformation sharing on Twitter", *arXiv preprint arXiv:2003.13907*.
Taylor D., *Code release: Supracentrality*, available at https://github.com/taylordr/Supracentrality.
Taylor D., Porter M.A, Mucha P.J. (2019), "Tunable eigenvector-based centralities for multiplex and temporal networks", *arXiv:1904.02059*.
Tremayne M. (2014), "Anatomy of protest in the digital era: A network analysis of Twitter and Occupy Wall Street", *Social Movement Studies*, 13 (1), pp. 110-126.
van Dijck J. (2011), "Tracing Twitter: The rise of a microblogging platform", *International Journal of Media & Cultural Politics*, 7 (3), pp. 333-348.
van Dijck J. (2013), *The culture of connectivity: A critical history of social media*, Oxford University Press, Oxford.
Yoon H.Y., Park H.W. (2014), "Strategies affecting Twitter-based networking pattern of South Korean politicians: Social network analysis and exponential random graph model", *Quality & Quantity*, 48 (1), pp. 409-423.
Zarocostas J. (2020), "How to fight an infodemic", *The Lancet*, 395 (10225), p. 676.
Ziems C., He B., Soni S., Kumar S. (2020), "Racism is a Virus: Anti-Asian Hate and Counterhate in Social Media during the COVID-19 Crisis", *arXiv preprint arXiv:2005.12423*.

9. Testing historical theories with SNA. Structure and evolution of a credit network

by Francesca Odella*, Cinzia Lorandini**

1. Introduction

When, in 997 the review of Erickson described the state of the art of social network approaches to historical data, she highlighted the potential as well as the requirements for performing a real structural analysis of historical phenomena. The prevalence of structure over attributes of social relations, she suggested, was to become the key concept to guide future researches, and the organization of historical datasets was required to overpass a descriptive aim and being targeted to explanation. The works of Gould (1991, 1996), Padgett and Ansell (1993) paved the way for a closer integration of social network methods and history and inaugurated a new perspective for testing hypothesis concerning social configurations of historical relationships (Franzosi & Mohr, 1997). The current research scenario, thanks to computational supports and sources digitalization, is flourishing and researchers are committed to expand the boundaries of SNA applications to great textual corpuses and longitudinal archive data (Morrissey, 2015). To escape a simply descriptivist approach, however, social networks analysis of historical cases must contemplate an accurate observation of the ancient contexts and accomplish reputable theoretical assumptions and interpretations of the study objects. As clearly stated by Gondal and Mc Lean (2013) "the meaning of relations, come to be patterned on the basis of identities acting across networks" that structural approach can bring to light and plausibly organize.

* Department of Sociology and Social Research, University of Trento, francesca.odella@unitn.it.

** Department of Economics and Management, University of Trento, cinzia.lorandini@unitn.it.

Against the backdrop of this methodological scenario, this research work aims at investigating the structure of lending relations in a pre-modern economy; the study focuses on the network established by a merchant family along approx. forty years (1747-1786) and scrutinizes two hypothesis concerning the social and economic mechanisms of early modern credit markets. The use of multiple archival sources (Raab, 2004), specifically, allowed us to explore in details the various social relations that affected the lending activity of the merchant family and the implications of social proximity in structuring economic relations between different social classes. Structural analysis and comparison of relations among different social groups involved in the credit network put in evidence, as well, the connections between notarial and non-notarial credit circuits and their mutual functionality for the lending activity.

2. Case study description

The study concerns a family commercial house – the Salvadori firm – that achieved considerable economic and social advancement in the eighteenth century in the Prince-Bishopric of Trento, a political entity in the central eastern Alps. Likewise to other merchants trading over long distances, they profited from excess liquidity to engage in lending, thus meeting the financial requirements of several types of borrowers – among them merchants, artisans, professionals, clergymen, widows as well as noblemen and institutions.

Our analysis aims at re-constructing the credit network of the family business and interpreting the relations among clients, patrons and other actors involved in the Salvadori's lending activity. The primary data source was the family credit registry (loans' ledger) which shows all the lending positions open in a period of more intense financial activity (1747 to 1786) for the merchant family. Loans in the registry were reported as "credit entries" and for each of them were found information concerning the identity of the borrower/s, length and amount of loans and interest rate, as well as mentions about warranties and negotiations, or repayment episodes. Moreover, if the loan was secured by a legal (notarized) document, the registry mentioned also the names of notaries, of the guarantees and other relevant actors involved in collateral activities related to the lending contract (such as payments transfer for petty personal expenses). To reconstruct the whole credit network of the Salvadori family we transcribed and recoded all the name citations that appear in the credit registry. We also supplemented the main source with other information from the Salvadori family history (Lorandini, 2006, 2015), and from other local historical archives[1]. These sources were relevant for identification of personal contacts and their collocation in terms of social position, intensity of acquaintance, professions, residence and business connections with the Salvadori. Cross-information about all the cited persons allowed us to establish different types of links among the lenders (Salvadori family) and the clients (an extract of the civil society of the time). The fine-grained and exhaustive information that we could retrieve about these actors convince us that the Salvadori business case can stand out for testing economic history hypothesis portrayed by recent theories.

Before the rise of modern banking, personal and commercial credit relations were influenced by social proximity (e.g. similar social class/profession) and regional conventions (local institutions); specifically, loans were regularly settled in the form of private agreements and fiduciary relations between the borrower and the lender (Carboni & Muzzarelli, 2014). Different social classes, therefore, had diverse opportunity to access credit and to invest or respond to financial fluctuations. Merchants, specifically, preferred not to resort to notaries for their business transactions and when they could, they avoided spending time and money for the public registration of contracts (Gelderblom *et al.*, 2018). The studies of the French credit markets by Philip Hoffman *et al.* (1999, 2000, 2019), however, documented the presence of a pervasive "shadow" credit market that pivoted on notaries. Results of extensive archival researches performed on the basis of Hoffman's *et al.* thesis established that in early modern societies notaries acted as brokers who matched lenders with borrowers and certified the borrowers' creditworthiness, thus contributing to expansion of credit. Hence, notaries did not limit themselves to drawing up and keeping legal records of lending contracts but performed also an intermediary function (Clemens & Reupke, 2009). Due to their privileged position, they knew who was in need or in excess of money and helped overcome the problem of asymmetric information by providing the lenders with information on the borrowers and their collaterals. Extensive empirical studies carried on in early modern rural Germany (Stark in Gestrich & Stark, 2015) account also that the intermediation of notaries was functional to borrowers for coping with the script and contents of the loan contract (guarantees, further agreements due to inheritance changes) and to lenders for securing large financial transactions.

[1] Archivio di Stato di Trento (State Archives in Trento), Archivio Salvadori (Salvadori Archives), vol. 734. The Salvadori business documents are conserved in the State Archives of Trento; an inventory project financed by Fondazione CARITRO is still in progress; hence we use the old archival references.

Following these relevant theories and research findings, the design of our study focused on tracing all non-notarial and notarial transactions that register a direct financial relation between the lender and the borrowers. The main units of investigation are the lending positions recorded in the family ledger: each position is related to a main client (borrower) and to a variety of subjects involved in the transaction as either writer of the deed (notary), guarantee or other financial role.

Accordingly, the data extracted from the Salvadori archive can be classified in two main types:

– information about the borrowers and other actors (roles, type of relations to the Salvadori – e.g. kin, agents, partners and business correspondents, and details about the social position);
– information about the lending positions and loans (amount, duration and type of credit, rates of interest).

We were able to detect 152 different lending positions for a quite limited period of time (1747-1767, with a last credit position opened in 1786) and to sum up more than three hundred of name citations (397 multiple citations), identifying 206 different actors, from local businessmen and affluent notables to small artisans and clergymen[2]. In the coding process we decided to insert also a taxonomy of the subjects which could be useful for comparing roles (e.g. subject involved at the origins of the loan and those involved in the subsequent transactions such as payers and receivers of credit for Salvadori family), and organized the data for network structural analysis (Alexander & Danowski, 1997). The raw data were then digitalized and coded according to standards suitable for the analysis with specific SNA software (Ucinet v 6.682, Borgatti, Everett & Freeman, 2002). The analysis was performed on the 152 by 206 matrix of personal relations (2-mode); for testing the hypothesis we used also 1-mode projections (206 x 206 matrices, with valued and binary data).

Table 1 reports the main characteristics of the actors registered inside the family ledger (by type/role in the transaction) and the frequency of citations for each type of actors[3]. We differentiated, in particular, between actors involved directly in the transaction (borrowers, notaries and guarantees), actors related to origins and developments of the loans, and actors in-

[2] Attribute variables of the cited actors include: social class (3 classes + 1 for institutions), role in the lending relation (main borrower, guarantee, notary, person involved in renegotiation, etc.), place of residence, amount (florins) and duration of the loan, and eventual presence of litigation about repayment.

[3] Multiple citations and multiple lending positions were present in the original data.

directly related nonetheless considered relevant by the ledger's writer. For the majority of the actors was also possible to distinguish the social class, on the bases of their professions and other information retrieved from the local archives.

Tab. 1 – Main characteristics of the actors and their frequency of citations

Type of subjects cited in the archive		*Frequency of citations*
Borrowers (person or institution)		152
Notary		63
Guarantee		16
Other types of subjects cited at the loan origin		65
Subject involved in the transaction (e.g. debt renegotiation/transfer)		68
Other person or institution cited during the loan duration		33
Total of citations (multiple name citations) by Social Class		
Social class of the cited actors and borrowers	*Cited actors*	*Borrowers*
Institutions	10	11
Upper class and nobles	37	42
Merchants, landowners, professionals	65	35
Artisans, petty traders or other	96	64

3. Approach and hypothesis formalization

Our approach is centred on the Salvadori's lending network as a whole unit of investigation: this means to analyse the characteristics of subjects in relation to a specific position in the financial transactions and their role in the global structure of relations (White, 1992). Then, we focus our analysis on two hypotheses: each one operationalizes assumptions concerning the historical development of early modern credit markets. The role of notaries, in particular, was investigated to support literature statements about their mediation role in providing access to credit for intermediate social classes.

Hp1. The lending business strategy of pre-modern merchants involved decisions based equally on social and economic conveniences. Creation and maintenance of the lending network was influenced both by social rules delimited by Salvadori's class position and by their commercial activity (organization and control of associates). To analyse the outcomes of the Salvadori business strategy we investigated the subnetworks generated by non-notarial credit connections versus the subnetwork generated by notary credit connections.

Hp2. Maintenance of the lending network in the local context required that the Salvadori family make available credit to clients of different social classes. The notary acted as referent/mediator and rule-guarantee for different social classes, and as indirect effect their third-party role results in opportunities to access credit for a larger segment of population, mainly belonging to middle-lower classes. To scrutinize the role of the notary we compared the density of relations among different groups of actors mentioned in the transactions and evaluate the role of notaries in relation to their possible bridging position.

4. The lending network

4.1. Social composition

The linkages between the credit positions and the identity of the cited persons in the archive allow us to reconstruct the complete credit network of the Salvadori family, which was constituted of 206 actors. Some of these actors tend to be frequently cited in the ledger, with different roles in the transaction. A more detailed distribution of actors (Table 2a), shows us that while merchants and people from lower classes where among the most frequently cited, half of the loans' positions (77 out of 152) were granted to borrowers from the higher classes (nobles and patricians) and to merchants or professionals. These figures support the interpretation that despite lending contracts commonly were arranged as peer-to-peer settlements, members of the lower class could also access them under specific conditions.

As one of the most important conditions was personal trust and reputation from social proximity, we additionally classified borrowers according to the intensity of their relations with the Salvadori family (Table 2a). Priority in intensity of relations (strong) was given to established business partners, commercial agents and extended family members (bonding relations). The other borrowers were listed according to their intensity of relations with the Salvadori based on frequency of business contacts with the Salvadori (e.g. renegotiation and successful closure of several credits). Social class and intensity of relations (strong/middle/weak social proximity) were then used for the testing procedures discussed in next sections.

Tab. 2 – Actors and borrowers by frequency of citations, social class and social proximity with the Salvadori family

a) N. of citations in the ledger	*Institutions*	*Noble class and patricians*	*Merchants and professionals*	*Lower classes*
1	6	20	30	75
2		6	13	15
3	3	8	5	3
4			4	2
5			4	1
6 and more	1	3	9	
Actors	10	38	63	95
Borrowers	11	42	35	64
b) Social proximity (intensity)				
Strong		8	9	3
Middle		8	19	21
Weak		26	7	40

The Salvadori made scant recourse to intermediaries for loans granted to members of the same class or to noble and patrician families which belonged to the local social and political élite (Table 3). Likewise, trust and reputation deriving either by kin or regular contacts – as in the case of business correspondents who already had a current account in the firm's main ledger – increased the preference for non-notarized loans. By contrast, notaries intervened much more frequently when loans were granted to the lower class, and only in very few cases member of this class could count on trust derived from intensity of contacts. This highlights that the Salvadoris' lending activity can be properly understood only within the framework of the overall economic and social strategies of the family.

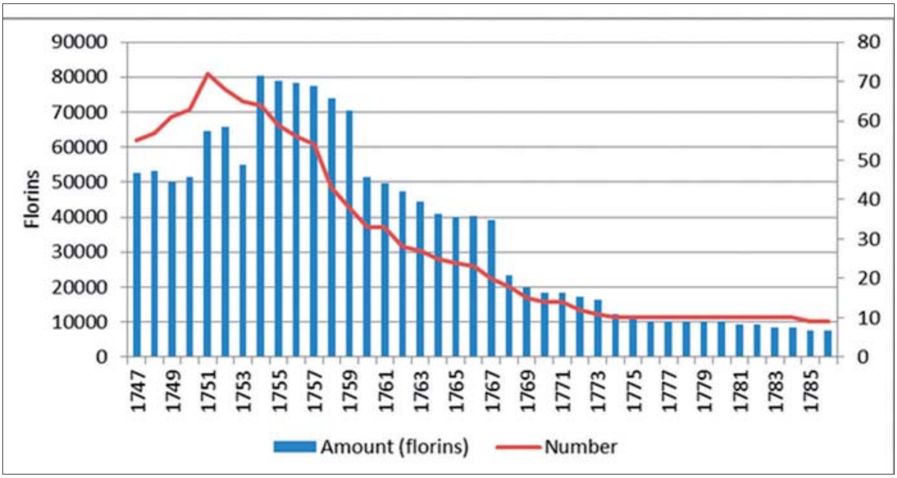
Tab. 3 – Notarial and non-notarial loans for borrowers' social class

	Institutions	*Noble class and patricians*	*Merchants and professionals*	*Lower classes*
Non-notarial loans	5	34	33	22
Notarial loans	6	8	2	42
Borrowers	11	42	35	64
% notarial loans granted to weak relations		26,9%	28,6%	75%

4.2. Financial activity

Analysis of the loan ledger tells us that the Salvadori were dedicated to large loans. At a time when specialized banking had still to take root, the Salvadori belonged to the restricted group of wealthy merchant families who addressed borrowers with substantial financial requirements, flanking local charitable and assistance institutions, which applied the same interest rates. The loan size ranged from 41 florins to around 14,000 florins, averaging about 1,100 florins (Fig. 1). Non-notarial loans (94 on a total of 152) were actually private transactions in the form of a simple person to person contract (48 positions), or a more articulated 'socially mediated' transaction (46 positions), which usually involved one or two persons together with the borrower. They were also generally larger than notarial ones, the higher average of which was only due to investments (equity and deposit) in a company promoted by the urban government, which was included among institutions[4].

Fig. 1 – Loans outstanding: number and value (1747-1786)



The peak of financial activity, with higher investments, is concentrated in a short period (1755-1759), when the merchant family had reached a

[4] In two cases, deposits represented actually an investment of risk capital in limited partnerships promoted by the town government, which was remunerated with a dividend. Both were silk firms promoted by the town government with participation of patricians and prominent merchants.

prominent role in the local economy. Loans issued by the Salvadori lasted even several decades and credit positions were frequently rolled over, as emerges from the substantial difference between expected and actual duration (Table 4). Almost 60 per cent of the loans with a specified duration were to last up until 2 years, but less than 30 percent expired within that date. The planned and actual duration were both higher for notarial loans compared to non-notarial ones, but in both cases the maximum length rarely exceeds 50 years.

Finally, the interest rates. In the eighteenth century usury restrictions took the form of an interest ceiling that was determined by local authorities. In the Prince-Bishopric of Trento, different interest rates were allowed based on the type of contract and the parties involved. The Salvadori family, however, granted only few loans at 6 per cent (the higher rate), and these exclusively to merchants active in the remunerative trade of silk; while for most of the lending contracts they granted a 5 percent interest and in some cases 4,5 percent or less for some loans granted to relatives, or loans deriving from selling land property. These findings confirm as argued by Hoffman *et al.* (2000), that the lender's "decision to make a loan may […] depend less on interest rates than on personal information about borrowers and extra-market relationships with them".

Tab. 4 – Loans' duration by type and amount

Loans		*Planned duration*		*Actual duration*		*Size (fl.)*	
	N	*Av.*	*Std.D*	*Av.*	*Std.D*	*Av.*	*Std.D*
Not-notarial loans							
	94	2.0	1,7	9.1	10.4	992	1,608
Notarial loans							
	58	3.3	1,9	14.8	12.4	1,295	2,746
Total	152	2.6	1,9	11.2	11.5	1,108	2,112

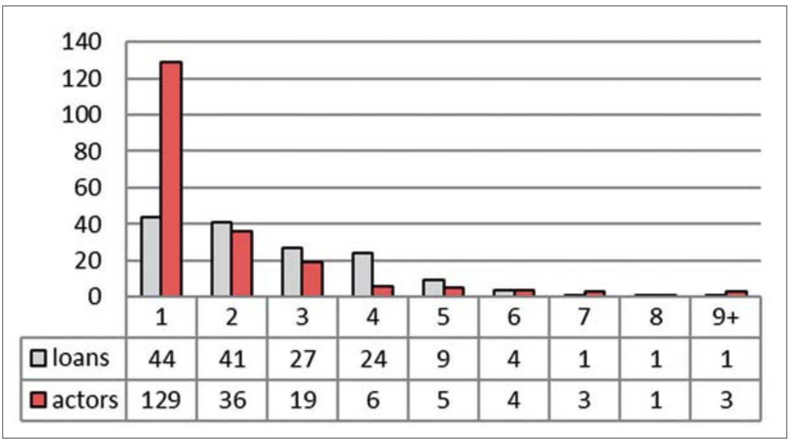
4.3. Structural analysis

The analysis of the credit network focuses on its structural features, dimensions and type of components, centrality and connectivity related to the presence of multiple links among actors, and scrutinizes the eventual presence of overlapping sub-networks.

The two-mode network of the actors by loans relations (Figure 2 for degree distributions for the actors) is characterized by a density overall of

0.014, with 47 different components (each represents a subgroup of credits and actors interconnected)[5]. This is congruent with the fact that a large part of loans was granted directly to individuals (44) with no other's presence (no notary or co-borrower). When loans involved multiple borrowers, they resulted from complex credit contracts (such as those linked to inheritances) or were renegotiated and sold to other lenders. Some actors, however, were involved directly and indirectly in multiple loans and hence mentioned several times with different affiliation and roles (borrower, payer, etc.).

Fig. 2 – Degree distribution for actors and loans



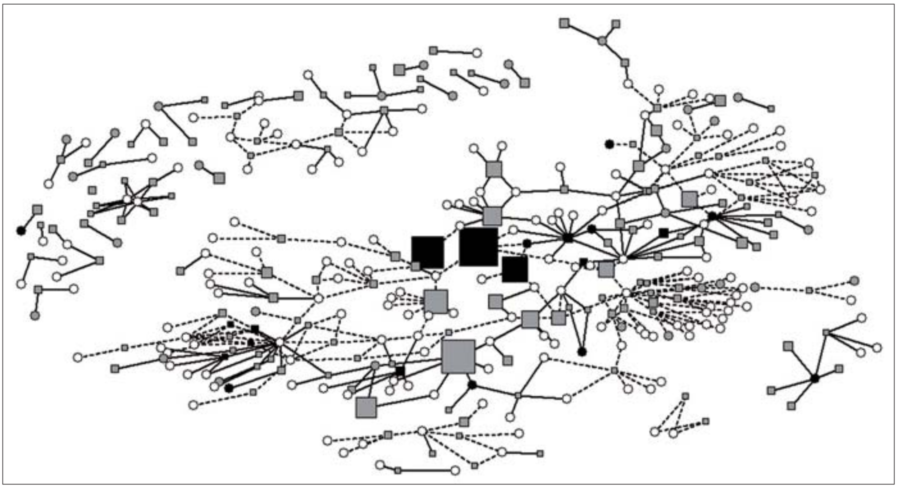
The visualization of the whole Salvadori credit network (Fig. 3) puts further in evidence the presence of several separate credit operations carried on with borrowers (a single person or same family members) having no specific social or business relation with the Salvadori business partners or other clients[6]. These transactions, moreover, usually did not involve large amounts of money (with the exception of a short term loan in 1750) even when the contract involved established commercial partners and extended family members. We can thus presume that the social relations involved in these loans had no special role for the dynamics of the lending activity.

[5] AVD for actors 1.93 (SD 1.88) and range 1-15; AVD for loans 2.61 (SD 1.59) and range 1-10. If we exclude the actors that are mentioned just one time the AVD rises to 3,48, meaning that some actors had multiple roles in relation with the 152 loans transactions. Loans with more than one actor involved have AVD of 3.26.

[6] Measures for the 2-mode network with all actors (206X152): 397 lines, density 0.013, AV. Deg. 2.612 and SD 0.112, AVD 7.716, Frag. 0.533, Trans. 0.602.

Instead, at the centre of the whole network is positioned a large group of loans characterized by larger amount of money and high occurrence of links between the borrowers and the other actors. This group of loans represents the major part of the Salvadori moneylending activity, related to important commercial transactions, involving family members and business partners. Since the structure of the network is an outcome of different roles in the transactions, to evaluate the interconnectedness of the relations we measured the density in the core of the 2-mode network, first for the complete network with all actors and loans (0.025), and then separately for the notarial loans (0.021) and the non-notarial loans (0.015). The results confirm that notaries are somehow involved in the core activity and may have played a role in relating different financial operations/ mediating contracts, or in providing new clients for long-term financial investments of the Salvadori family.

Fig. 3 – The credit network of the Salvadori family[7]



[7] Graphic of the 2-mode network. Squares (sized by amount) represent loans and circles represent actors; nodes' colour is scaled by social proximity intensity; black colour stands for institutions. Dashed lines are notarial connections.

5. Testing the hypothesis on the lending network

5.1. Strategic organization of lending activity (Hp.1)

To test the first part of the hypothesis concerning the strategic management of the credits we created two separate sub-networks that contain respectively: a) the relations directly and indirectly produced by cross-references dealing only with non-notarial contracts between Salvadori and their clients; and b) the relations produced directly and indirectly by cross-references dealing only with notary acts signed by the borrowers.

Relying on previous literature suggestions (Clemens & Reupke, 2009), we assume that the notary played a bridging role eminently for the lower classes and small entrepreneurs and only in limited cases (large investments) the more affluent Salvadori's clients were involved in notary acts. As for the rest of the credit positions (approximately two thirds of the lending activity) loans were granted by the Salvadori through fiduciary contracts to small business correspondents, as well as to borrowers belonging to middle and upper classes.

Tab. 5 – Structural measures for type of subnetworks[8]

	Notarial loans	*Non-notarial loans*
Nodes	206	206
Ties	594	356
AV Degree	2,883	1.728
Density	0.014	0.008
Centralization (D)	0.074	0.070
N components	68	129
Fragmentation	0.846	0.937
Closure	0.631	0.710
Network Transitivity	0.221	0.449
AV Distance	4,320	3.649

The analysis of two distinguished subnetworks resulting from the selection procedure exemplifies two different types of generative linking mechanisms at the bases of credit markets and their referral mechanism (White, 2002). Connections deriving from notarial documents and legal procedures (inheritances, land and property investments) tend to create closed circuit of

[8] Values for actors by actors matrix, 1-mode network, binary.

concurrent roles (notaries are related to other notaries, guarantees to other guarantees): when reproduced through time this generative mechanism creates a series of recurrent links among some of the actors (revealed by higher density but lower transitivity indexes in the sub-network of notarized loans). On the other side, connections based on strict commercial and business-oriented relations are based on substitutive chain mechanisms (as business partners may change according to the nature of the deal) and tend to generate a structure with a more linear shape (indexes of this sub-network display in fact lower average degree and distance). Commercial credit transactions also rely on mediators that can play multiplex roles in a network and relate credit demand and supply at the right time for a profitable investment (Haggerty & Haggerty, 2013). The presence of notaries was thus important in supporting the Salvadori lending business, helping to match both financial and social conveniences.

Concluding, both generative mechanisms combined to produce the complete lending network; the co-occurrence of the two business strategies was also facilitated by bridging partners with a strong centrality position in the network (and among them we find a significant proportion of notaries)[9]. Subjects such as trade partners and fiduciary notaries presumably were also able to support the financial strategy of the Salvadori family and/or were involved in the financial and social implications of their moneylending activities.

5.2. The mediation role of the notary (Hp.2)

According to the hypothesis that notaries distinct public role endorsed access to credit for emergent social classes, our analysis in this paragraph consists in a quantitative assessment of the role of notaries in the structure of the Salvadori lending network.

First, to analyse the contribution of notaries on the structure of the lending business network, we perform an analysis of sub-networks extracted from the complete lending network. As illustrated before, loans associated to a notary document tend to be overrepresented in the main component of the lending network (resulting in a higher core density) and are also significantly related to larger credit amounts. Therefore, to test the hypothesis of the bridging role of the notary in the business network we first juxtapose typical

[9] Analysis performed on individual nodes (1-mode network projections) showed that notaries tend to score the highest values for several centrality measures (degree, betweenness and 2-local eigenvector centrality), together with business agents of the Salvadori and few patricians.

forms of credit situations in the Salvadori registry that involved an increasing influence of the notary over the actors involved in the credit transaction.

We thus compared (Table 6) the structural characteristics (overall density and other measures of cohesion) for the sub-networks generated by two types of relations: a) links among the borrowers and all the other actors, including those actors indirectly related to the transactions (receivers of credit from the Salvadori, other actors mentioned in the registry); b) links among actors that are directly related to the credit document (only borrowers, notaries and guarantees). To control for the hypothesis of bridging we also computed the structural values for the whole network without the actors with a notary role (Control Case).

Specifically, direct relations generate a sub-network (B) that is composed of fragmented independent units that display no interconnectedness; on the opposite side, the presence of both direct and indirect relations produces a sub-network (A) characterized by less components and higher centrality and closure, hence a context where actors interrelate and associate among them.

Indirect relations are crucial for establishing (and maintaining) an efficient business system (Mc Lean & Padgett, 1997) and besides notaries, several people – acting as guarantors, sellers or buyers of receivables, payers on behalf of the original borrower, partners of the borrowing company, and so on – played a role in the functioning of the credit market by channelling information and facilitating contacts with other actors.

Tab. 6 – Comparison between networks (hypothesis on the role of the notary)[10]

	Network A	*Network B*	*Network C*
	Direct and indirect relations	*Only direct relations with the contract*	*Control case*
Nodes	206	206	206
AV Degree	4,165	0,544	2,883
Centralization (D)	0,142	0,061	0,074
N components	28	154	68
Fragmentation	0,486	0,990	0,846
Closure	0,484	0,081	0,631
Average Distance	3,995	1,857	4,320

To evaluate the specific bridging role of notaries in the Salvadori lending activity we compare the control case (a sub-network without notaries) with sub-network A. Substantive variations for the average degree, as well as for

10 Values for actors by actors matrix, 1-mode network, binary.

closure and fragmentation indexes between the two sub-networks testify that notaries may have contributed to link actors other than borrowers (as the average distance increases in the control network), and to crossing social class boundaries, especially when property and inheritance transactions were involved.

Second, to evaluate the role of notaries in bridging among social circles we focus on the loans supported by a notary act. One third of the lending positions (58 out of 152, two positions having multiple acts) are supported or incorporated in a notary manuscript, which frequently includes supplementary information about the transaction (borrowers' assets, his/her family situation and other legal relevant information). The ultimate reason to provide such detailed financial profile of clients may be related to the association of social class and reputation from social proximity. This is supported by the fact that in our data notarized loans are concomitant more with transactions involving actors that are considered less trustful or have no previous commercial experience with the Salvadori.

To put in evidence differences among types of contracts (and distinct associative patterns among actors) we consider the density of relations among homogenous groups of actors[11]. Groups are constituted by subjects that share the same level of intensity of relations with the Salvadori (strong/middle/weak social proximity). Then, to evaluate the economic function of notarized contracts we compare the density of relations that are present in the two previous subnetworks (notarized and non-notarized loans of Table 5) with a third sub-networks that contains only the actors involved in the financial transactions (borrowers, guarantees, payers/receivers of credit and the notary). This sub-network – defined as Network T – is supposedly less influenced by subjective elements (ex. ledger's transcription habits) than the complete network – here inserted for completion – and may suit better for testing our hypothesis (Table 7).

The density value previously calculated for the complete network (all types of contracts) shows that among actors with intense social proximity to the Salvadori – either by familiarity, patronage, or business linkages – there is strong interconnectivity (density 0.056); while groups of actors that report low or middle intensity of relations with the lenders show a lower internal density (0.022 and 0.012).

Since the internal cohesion of each group is also related to the type of links among actors (sub-network of contracts and subnetwork of notarized loans), comparison of group density for network T, created by the removal of "additional" linkages improves our argument. Results show that actors with

[11] Institutions were not included in this part of the structural analysis.

a strong social proximity with the Salvadori are associated to credit circuits regardless of the form of contract that was issued for lending (Table 7a). On the contrary, those actors that did not possess trustfulness or financial reputation were associated more frequently to notarized transactions[12]. To understand how social proximity is related to loan types (private agreement/contract vs notarized loan), we additionally computed the values of the densities of the three groups controlling for the effect of social class (Table 7b). Comparing the complete network to sub-network T, the closure effect that was determined by the presence of different social classes is evident for the actors sharing high social proximity with the Salvadori, and less obvious for those actors having a middle intensity of relations.

Tab. 7 – Comparison between networks (group density by intensity of relations)

	Weak	*Middle*	*Strong*
a) Social proximity			
Contract	0,009	0,005	0,028
Notarized loans	0,013	0,007	0,028
Network T	0,014	0,005	0,028
Complete network	0,022	0,012	0,056
b) Social proximity with control for social classes			
Contract	0,012	0,007	0,014
Notarized loans	0,013	0,005	0,018
Network T	0,015	0,009	0,020
Complete Network	0,023	0,013	0,032

Concluding, the overall structural analysis confirms both hypotheses with supplementary details. Social class of the borrowers had ultimately an impact on the business strategy of the Salvadori. Social class specifically increased the divide among those who did not require documented credentials to access loans (actors with middle and higher proximity) and those necessitating of the notary intermediation. Nevertheless, the Salvadori were judicious in offering differentiated types of loans (notarized and private contract) to clients of different social classes, being intensity of interactions or social proximity an important requirement to stipulate secure credit transactions (Hypothesis 1). The presence of notaries – presence not necessarily related to the legal act – enhanced also traditional exchanges, usually stipulated among

[12] As the Salvadori ledger entries testify, notary loans were also a form of "social credentials" that were used in the lending market to support the borrowers, as well as to re-negotiate the credit in the following years.

subjects belonging to a similar social milieu, and hence based strictly on professional reputation and assets. As locally based professionals, notaries proved particularly crucial in establishing connections among their clients of different social classes and ensure the legal conditions of the transaction, as well as its economic accomplishment (Hypothesis 2)[13].

6. Final comments

Which factors drove a lender to resort to a notarized loan? In the early modern economy this decision was determined mostly by the lender's characteristics and the characteristics of the prospecting clients. On one side, membership in the same social group was a priority factor that enhanced trust between the lender and the borrower and private dealings were the usual choice for loans to relatives, friends and business associates. On the other hand, the higher the uncertainty and the lower the trustworthiness of the counterparty, the higher the incentive to register with a notary. And even if according to social conventions of the time it was not mandatory to hold lending contracts with a notary act, notaries established a framework for trust and respect of rules among subjects belonging to different social classes (Levy, 2010). Their presence – not necessarily related to the legal act – enhanced traditional exchanges, usually stipulated among subjects belonging to a similar social milieu.

From this standpoint, our intent to show how Social Network Analysis can be applied to support hypothesis concerning the structural configuration of lending relationships has been partially fulfilled. As we have illustrated, strict financial relations, in fact, may induce to underestimate the combined effect of socio-economic institutes and social classes in structuring dissimilar paths to access credit. Loans that are not supplemented by a notary act were more frequently arranged among customary business partners or trustful clients, both sharing social proximity with the lender; while for lower social classes and specifically artisans or small landowners, notary acts might constitute supplementary references about their assets (ex. inheritance of land properties, investment plans) and economic reputation.

Specifically, in the peripheral local economy of the Prince-Bishopric of Trento, where our case is situated, strategic organization of lending required an accurate selection of patrons (nobles and patricians), institutional

[13] On the topic see the analysis of the letters of credit by Padgett and McLean (2011) and Gondal and McLean (2013) in Renaissance credit market.

representatives and prominent commercial partners, as well as arm-length trading with lower classes and artisans that could enhance and provide stability to the main commercial business. Focus on the local context is not secondary to our results: as previous research has documented, explanations of credit business circuits rely also on the sources of information about the financial activity and the methodology that is adopted to investigate historical cases (Reupke, 2015; Muldrew, 1998). The assessment of the hypothesis concerning strategic organization of the lending activity and the mediation role of notaries show in fact that new historical documentation can lead to original perspectives of the borrower-lender relation. Family archives and small case studies, such as litigation records, despite being more fragmentary and circumscribed to local economies, may in this sense provide a novel source for testing determined hypotesis via direct historical sources.

References

Alexander M.C., Danowsky J.A. (1990), "Analysis of an Ancient Network: Personal Communication and the Study of Social Structure in a Past Society", *Social Networks*, 12, 4, pp. 313-335.

Borgatti S., Everett M., Freeman L. (2002), *Ucinet for Windows. Software for Social Network Analysis*, Analytic Technologies, Harvard.

Carboni M., Muzzarelli M.G. (a cura di) (2014), *Reti di credito: circuiti informali, impropri, nascosti (secoli XIII-XIX)*, il Mulino, Bologna.

Clemens G.B., Reupke D. (2009), "Der Notar als Broker: Das Management des privaten Kreditmarktes", *Zeitschrift für Verbraucher- und Privat-Insolvenzrecht*, 8-12, pp. 16-22.

Erickson B.H. (1997), "Social Networks and History: a review essay", *Historical Methods*, 30, 3, pp. 149-157.

Fontaine L. (2014), *The Moral Economy. Poverty, Credit, and Trust in Early Modern Europe*, Cambridge University Press, Cambridge.

Franzosi R., Mohr J.W. (1997), "New Directions in Formalization and Historical Analysis", *Theory and Society*, 26 (Special Double Issue on New Directions in Formalization and Historical Analysis), 2/3, pp. 133-160.

Gelderblom O., Hup M., Jonker J. (2018), "Public Functions, Private Markets: Credit Registration by Aldermen and Notaries in the Low Countries, 1500-1800", in M. Lorenzini, D'Maris Coffman, C. Lorandini (*eds.*), *Financing in Europe: Evolution, Coexistence and Complementarity of Lending Practices from the Middle Ages to Modern Times*, Palgrave Macmillan, London.

Gestrich A., Stark M. (*eds.*) (2015), *Debtors, Creditors and Their Networks. Social Dimensions of Monetary Dependence from the Seventh to the Twentieth Century*, German Historical Institute London, Bulletin Supplement 3.
Gondal N., McLean P. (2013), “Linking tie-meaning with network structure: variable connotations of personal lending in a multiple-network ecology”, *Poetics*, 41, pp. 122-150.
Gould R.V. (1991), “Multiple networks and mobilization in the Paris Commune, 1871”, *American Sociological Review*, 56, pp. 716-729.
Haggerty J., Haggerty S. (2013), “Visual Analytics for large-scale actor networks. A case study of Liverpool, 1750-1810”, in M. Casson, N. Haschimzade (*eds.*), *Large Databases in Economic History*, Routledge, London.
Hoffman P.T., Postel-Vinay G., Rosenthal J.-L. (1999), “Information and Economic History: How the Credit Market in Old Regime Paris Forces Us to Rethink the Transformation to Capitalism”, *American Historical Review*, 104, pp. 69-94.
Hoffman P.T., Postel-Vinay G., Rosenthal J.-L. (2000), *Priceless Markets. The Political Economy of Credit in Paris, 1660-1870*, University of Chicago Press, Chicago.
Hoffman P.T., Postel-Vinay G., Rosenthal J.-L. (2019), *Dark Matter Credit: The Development of Peer-to-Peer Lending and Banking in France*, Princeton University Press, Princeton.
Levy J. (2010), “Notaries and Credit Markets in Nineteenth Century Mexico”, *Business History Review*, 84, pp. 459-478.
Lorandini C. (2006), *Famiglia e impresa. I Salvadori di Trento nei secoli XVII e XVIII*, il Mulino, Bologna.
Lorandini C. (2015), “Looking beyond the Buddenbrooks Syndrome: The Salvadori Firm of Trento, 1660s-1880s”, *Business History*, 57, pp. 1005-1019.
Lorenzini M. (2018), “Notarial credit markets in eighteenth-century northern Italy: Trento and Rovereto compared”, in M. Lorenzini, C. Lorandini, D’Maris Coffman (*eds.*), *Financing in Europe: Evolution, Coexistence and Complementarity of Lending Practices from the Middle Ages to Modern Times*, Palgrave Macmillan, London.
McLean P.D., Padgett J.F. (1997), “Was Florence a Perfectly Competitive Market? Transactions Evidence from the Renaissance”, *Theory and Society*, 26, 2/3, pp. 209-244.
Morrissey R.M. (2015), “Archives of Connections”, *Historical Methods*, 48, 3, pp. 67-79.
Muldrew C. (1998), *The Economy of Obligation. The Culture of Credit and Social Relations in Early Modern England*, Macmillan, Basingstoke-London.
Padgett J.F., Ansell C.K. (1993), “Robust Action and the rise of the Medici, 1400-1434”, *American Journal of Sociology*, 98, pp. 1259-1319.
Padgett J.F., McLean P.D. (2011, March), “Economic Credit in Renaissance Florence”, *The Journal of Modern History*, 83, pp. 1-47.
Raab J. (2004), *From Words to Numbers to Graphics: A Suggestion on How to Analyze Inter-Organizational Networks Using Archival Sources*, paper presented at XXIV Sunbelt Conference, Porteroz SLO.

Reupke D. (2015), "Credit Markets in the Nineteenth-Century Countryside: A Comparative Study in a Rural Border Region," in A. Gestrich, M. Stark (*eds.*), *Debtors, Creditors and Their Networks. Social Dimensions of Monetary Dependence from the Seventh to the Twentieth Century*, German Historical Institute London, Bulletin Supplement 3, pp. 119-143.

Stark M. (2015), "Networks of Lenders and Borrowers: A Rural Credit Market in the Nineteenth Century", in A. Gestrich, M. Stark (*eds.*), *Debtors, Creditors and Their Networks. Social Dimensions of Monetary Dependence from the Seventh to the Twentieth Century*, German Historical Institute London, Bulletin Supplement 3, pp. 99-118.

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

White H.C. (1992), "Cases are for identity, for explanation, or for control", in C.C. Ragin, H.S. Becker (*eds.*), *What is a Case? Exploring the Foundations of Social Inquiry*, Cambridge University Press, Cambridge.

White H.C. (2002), *Markets from networks: Socioeconomic models of production*, Princeton University Press, Princeton.

10. Innovative welfare networks. Ego-network analysis of innovative startups "with social vocation" (SIaVS) in Piemonte and Campania

by Massimo Del Forno*, Marco Di Gregorio**

1. The Network Europe and the challenge for social impact

The background for this paper is the social impact measurement challenge, conceived in the European Union as leverage to involve and steer the private sector in addressing the welfare crisis. We focus on a particular Italian initiative concerning the building of an innovative welfare network around a new kind of "social" enterprise, the *innovative startup with a social vocation* (SIaVS). Since 2015, SIaVSs are the first experimentation of a process of institution-building to apply the requirement of social impact measurement on private companies, according to the principle of subsidiarity. We call "the norm of social impact" the set of ideas, methods, and practices which regulate the activities of social impact-driven enterprises. Despite a hopeful launch and some rhetoric, attention on SIaVSs lowered soon. The debate about the social impact measurement otherwise is increasingly in the spotlight while the norm of social impact is being extended to other kinds of business. Observing what happens around the SIaVSs allows us to identify some factors which favor or hinder the building *of innovative welfare networks* "from below" and the assumption of shared responsibilities among the actors of welfare policies. This paper was conceived as an early step in a wider research effort to understand how the social impact-driven startuppers perceive themselves in their networks and whether and how they relate with organizations of different nature sharing ideas and responsibilities. The par-

* Department of Political and Social Studies, University of Salerno, delforno@unisa.it.

** Department of Political and Social Studies, University of Salerno, mardigre@unisa.it.

The essay was designed by both authors. Massimo Del Forno edited §§ 1, 2. Marco di Gregorio edited §§ 3 and 4. The conclusion comes from a shared reflection.

ticular context of Italy also requires taking into account regional disparities and the structural gap between the North and the South.

Since this is an ongoing process, it is not possible to control any hypotheses of change. Nevertheless, making use of ego-network analysis in a qualitative approach to social research, we can check if there are some conditions for a change in the way of social impact-driven startuppers interpreter subsidiarity relations and social responsibilities. Before showing some findings from field research, it is useful to present in broad outline the background in which the issues of subsidiarity and social impact measurement lie that is the constituent process of the European Union and its rapid changes.

1.1. The first turn: the "Network Europe"

The first turning point dates back to 1992. Maastricht Treaty profoundly modifies the orientations to European integration, after dismissing of the old community myth that had inflamed the federalist vulgate. Carried by the wind of economic globalization, a new vision of *smart* and *mobile* Europe is emerging (Prodi, 2000; Castells, 2000a; Jansen & Richardson, 2004; Delanty & Rumford, 2005). A new form, which many call "Network Europe", was born. It is "a networked polity able to stake its claim in a networked and globalizing world", a fluid and flexible dispositive to promote the growth both of social and market exchanges (Delanty, 2005, p. 121). Network Europe is not a simple reaction to economic globalization but its most advanced expression. Network Europe and economic globalization share the same strategic objectives: the wellbeing culture, the globalization of markets, and advanced communication (Castells, 2000b). The network metaphor made any intergovernmental hypothesis inadequate, forcing Europe to renounce to become a new superpower or an extra-large nation and to pursue the experiment for the stateless government, working for agreements time and time again (Giddens, 2007). According to these assumptions, it is necessary to overcome the formal constitution of binding laws and norms, and re- thinking Europe as a polycentric structure, driven by a group of *inter pares*, with the aim of ensuring the economic and social cohesion of the Member States of the European Union. Any center is a node in a *competition-oriented space economy* (Richardson & Jensen, 2000). This is a relevant point to our research. This idea of Europe was methodically implemented using both the principle of subsidiarity (vertical and horizontal) and the new digital knowledge and technologies in order to stimulate the creation of a system of long networks connected with proximity networks. At the level of practices, these changes were associated with the order-word "evaluation". Meritocratic rules and procedures are established along with the development of many indicators to assess the performance of public policies and business operations to drive fundings across Europe. New organizations are needed to manage assessments, operations, and resources. A new organizational field is born.

1.2. The second turn: the European challenge

For a well-known rule that does not seem to admit exceptions, the emphasis on the means (efficiency) produces effects on the chain of ends (effectiveness). Primarily, the problems of the Network Europe emerged in the welfare arena. Despite the Lisbon Strategy – which would have wanted to make Europe "the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion" – even before the 2008 crisis, we were witnessing an increase in health and welfare costs, rising unemployment, and the widening of the gap between rich and poor. These problems become cross-cutting in Europe, along with the financial crisis, an increase in migration flows, demographic change, and the awareness of "global risks". Europeans agreed that in the balance between the economy and society, more weight should be lent to the social. Investments, therefore, must be bound to the evaluation of both programs and projects, to the point of making it mandatory to measure their economic and social impact on the territories.

In the last decade, the European Commission, together with OECD, the World Bank, and United Nations, determined a change in the social enterprise's world by dictating new rules and procedures to participate and access funds. Inspired by the themes of social economy, from one hand, and by the venture philanthropy, from the other, these new guidelines bind enterprises and investors to precise commitments: 1) undertaking to general interest; 2) measuring the social impact; 3) *being meritocratic*. Founds and incentives must reward virtuous and deserving enterprises that *build evidence* of added-value production. At the end of the *value chain*, the "social impact" shall address social inclusion, wellbeing, and sustainable development. Social responsibility and accountability concern both the economic and social effects of programs or activities. The emphasis on the "social", however, should not be misleading as it does not require any radical change in the economic order.

1.3. The circular subsidiary: a new cultural turn?

In broad outline, the issue of social impact in Europe has been developed in the mainstream political-economic vision and addressed to the new frontiers of the financial capitalism, the *impact investing*, through the innovations brought by the venture philanthropy and *philanthrocapitalism* (Bishop & Green, 2008). Even if, alongside the financial bottom line, the social and the environmental bottom lines have been added, the concept of "impact" remains a variable dependent on the stability of financial markets and the capitalist economy. The guiding idea is the social and environmental crisis must not slow down the economic development, but on the contrary, risks and welfare problems must be seen as opportunities for stimulating entrepreneurship and doing business, with the support of the state, the international and local public authorities, the philanthropy, and the for-profit capital altogether. In this perspective on welfare, which from the political point of view originates in the Third Way, the notion of social impact, rather than guaranteeing the centrality of people's well-being and health in the policy making, appears as an *effect-instrument* of the system centered on capital. In this process, public authorities and business organizations strictly preserve their respective spheres of competences, according to the vertical dimension of the principle of subsidiarity. However, this process seems to leave out the organized civil society, which is called at most as a residual function (precisely as Third Sector) to intervene in the spaces left uncovered by the State and the Market or to deal with their failures[1]. For this very reason, the emerging paradigm of Civil Economy assumes absolute importance to us. While the mainstream Political Economy faces welfare problems by classical focusing on the delegation of authority and responsibility to maximize the efficiency – or the use of resources which are scarce by definition –, Civil Economy works for the evolution of the idea of welfare towards the *civil welfare* (Zamagni, 2016), based on the *circular subsidiarity*. According to this principle, public authorities, market and business community organizations, and the organized civil society must share resources and responsibilities to address

[1] Neoliberal thinking looks to the civil society organizations as a safety net to ensure minimum levels of social services to the fragile people otherwise left behind by the dismantling of the welfare state (Zamagni, 2018, p. 18; Mazzuccato, 2013). With the *impact investing*, the for-profit business seeks a way to fill this niche of market. The criterion of "additionality" demands that impact investments should be directed precisely to undercapitalized areas where the state and the traditional investors back out, producing both financial and social returns (Calderini, 2019, p. 5). In the handover of care tasks from the welfare state to the social impact market, the proposals of civil society risk going unheeded, unless social impact orientation also means public participation in social planning and decisions on matters of general interest.

social problems without delegating, as a real community where the different functions of authority, economic rationality, and sociality give rise co-programming and co-projecting on territory (Moro, 2009). The new paradigm could have great transformative potential by its ability to create complex networks and circular links between people. To Civil economists, social impact means to bring change in the people's life by "civilizing" the market economy, while social impact assessment and measurement should be the leverage "to put the economy at the service of people" rather than the other way round, which translates into putting the people at the center of the "civil enterprises" activities.

1.4. Italy and the challenge of the social impact

Italy responds to the European challenge by offering social impact-oriented entrepreneurs support, tax breaks, or even just a qualification to add to the company name in exchange for their commitment to measuring the social impact generated. Having been a part of discussion panels set up by the government, Civil economists had a key role in the development of the norm of social impact in Italy. Nevertheless, due to the historical circumstances and the variety of positions on the field, the norm and its associated guidelines present many inconsistencies and unclear points. For the purpose of our research, this situation is very interesting because it leaves ample room for subjective interpretations for "social impact objectives" and the role of the enterprises in the innovative welfare network.

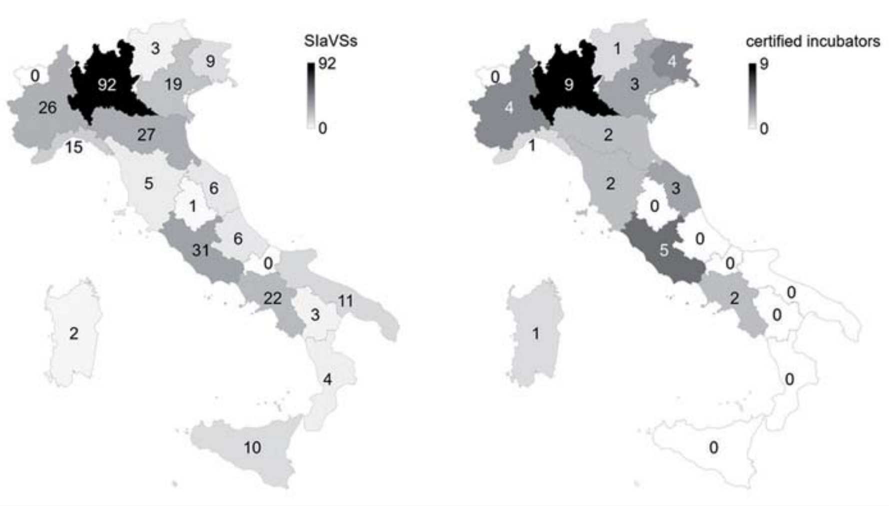
As a consequence of the application of the norm, SIaVSs have a *hybrid status*:
- startuppers must renounce at the distribution of profit for at least five years;
- they must work in the "general interest" sector, as well as third sector enterprises do (but they can either be third sector organizations or not);
- they must measure and report the social impact.

In return, startuppers receive fiscal incentives, preferential financing channels, training, and support by incubators and accelerators. Moreover, the norm confers a special certification to those incubators that meet a series of requirements and makes them privileged interlocutors in the network. Thus, in Italy, the organizational field born around the evaluation, assessment, and measurement of social impact is enriched with new elements, rules, and prizes. The organizational field is not a neutral network where you only learn technical notions, but it is also a political space in which charismatic visions of social impact can flow and grow. Startuppers can be fascinated by these ideas and build important ties, included personal, informal, and effective relationships. If the Civil Economy perspective on social impact prevails, we expect social impact-oriented startuppers to represent their networks as a community connecting different types of organizations and different institutional levels, where they share responsibilities, resources, projects, and activities for the well-being of persons on the territory.

2. The Italian network between SIaVSs and certified incubators

The register of companies shows 302 startups with "social vocation" from 1st January 2015 to the end of 2019; 215 are still active as SIaVS. We considered all the SIaVSs and certified incubators (from now on CIs) active between 01/01/2015 and 09/30/2019, for a total of 292 SIaVSs and 37 CI[2].

Fig. 1 – Distribution of SIaVS (on the left) and CIs (on the right) by region. Based on the head office, darker regions host more SIaVSs/CIs than the others



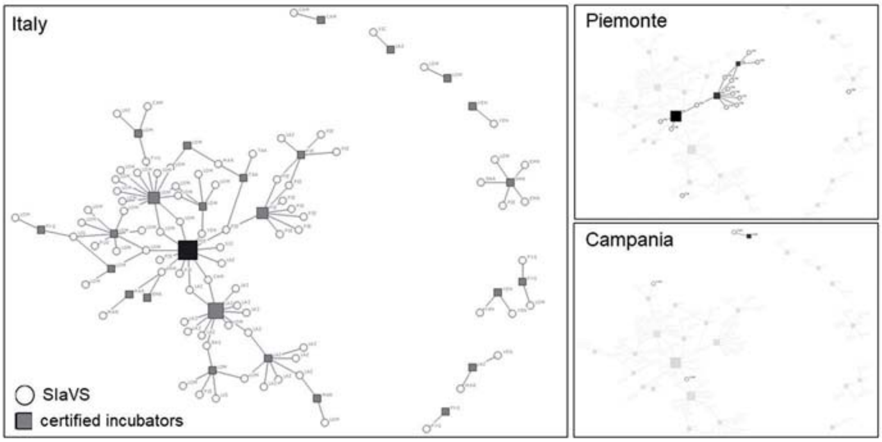
The distribution of both SIaVSs and CIs is very unbalanced in favor of Northern Italy with a strong concentration in Lombardia. The only region that hosts incubators in Southern Italy is Campania. Campania immediately follows Piemonte in the ranking by number of SIaVSs activated since 2015,

[2] An incubator from Lombardia obtained certification after 09/30/2019. Since it was very active even before, we included it in the analysis. The 38th, form Sardegna, is certified since 5/5/2020.

placing it as the first in the South. Fig. 2 shows the structure of the relationships between SIaVSs and CIs.

The largest of the 10 components in the Italian network includes about 80% of the nodes. The others are composed of a single CI and of one to five SIaVSs. The largest and darkest node is the CI with the highest *betweenness centrality*[3]. The comparison between Piemonte and Campania highlights how the companies from Piemonte are more active, more cohesive, and more central in the Italian network.

Fig. 2 – Two-mode network between CIs (grey squares) and SIaVSs (white circles)



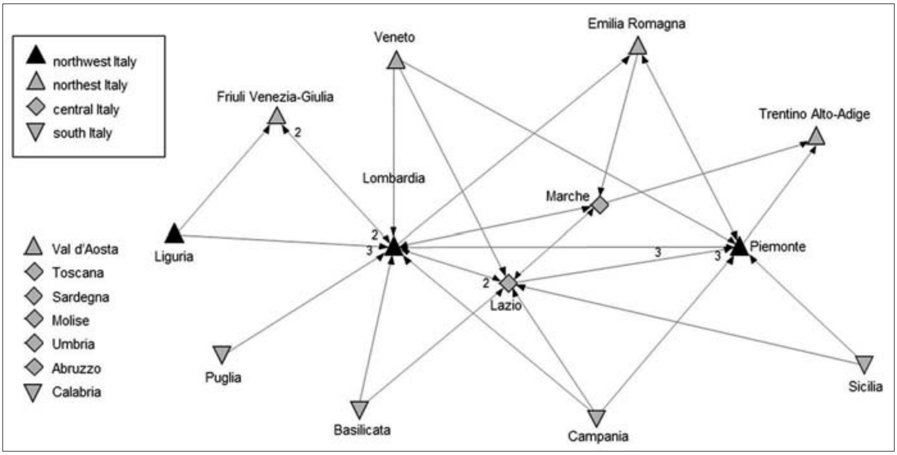
Nodes size is proportional to the betweenness centrality. Ties indicate the presence of at least one of these conditions: we found evidence of a formal relationship on the website or in official documents by the SIaVS or the incubator; economic relationship appears in a financial statement or in the register of companies. The two smaller boxes on the right highlight nodes from Piemonte (above) and Campania (below)

The structure of the ties between regions is an interesting issue because of the considerable autonomy of regional authorities to support SIaVSs and incubators. Likewise, incubators have different possibilities by region to offer services and resources for the growth of startups, networking with public and private institutions. In Fig. 3, nodes are the Italian regions, the weight and the direction of arcs indicate how many SIaVSs have been moving towards CIs in other regions.

[3] It is Socialfare, specialized incubator for social enterprises and important node of the Torino Social Impact's network. By degree, Socialfare has one point less than Impact Hub Milano (13 vs 12).

The graph highlights the profound inequalities between areas of the country. Except for Val d'Aosta, there are no northern regions among the isolated nodes. No SIaVS moves to the CIs in Campania (the only two in the Southern Italy). The relations between SIaVSs from the South and CIs can be counted on two hands.

Fig. 3 – Mobility of SIaVSs between Italian regions. Nodes are the regions



Arcs illustrate the movement of one or more SIaVSs from the region that hosts its/their head office towards certified incubators located in another region

In Italy, 28% of SIaVSs have formal ties with CIs. On a regional basis, only in Lazio and Piemonte it exceeds 50%. In Campania, 3 out of 22 SIaVSs go to CIs. Except for Lombardia, the differences between regions in absolute terms are small. The case of Piemonte stands out for the higher proportion of SIaVSs which are linked to an incubator (54%) and for the attractiveness of its CIs measured with the in-strength of ties[4]. This situation could be explained by the presence of a shared plan in a network of organizations called "Torino Social Impact" to make the city of Turin the European capital of social impact. The phenomenon should be observed over time to understand the effectiveness of the plan in sustaining *social impact-oriented* societies, attracting funds, promoting virtuous relationships, and driving social change. All of this noticeably is lacking in the South.

[4] With only 4 CIs, Piemonte attracts 10 SIAVSs from other regions. Lombardia attracts 12 SIAVSs but having 9 CIs.

Tab. 1 – Relationships between SIaVSs and certified incubators by region

	N. SIaVSs	*N. SIaVSs linked to CI*	*Ratio of SIaVSs linked to CIs on SIaVSs*	*N. CIs*	*N. CIs linked to SIaVSs*	*Reflexiveties*	Deg^{in}	Str^{in}	Deg^{out}	Str^{out}
Lombardia	92	28	.30	9	8	28	9	12	5	9
Lazio	31	16	.52	5	4	14	6	7	3	5
Emilia Rom.	27	4	.15	2	2	4	2	2	2	2
Piemonte	26	14	.54	4	3	15	6	10	3	3
Campania	22	3	.14	2	1	1	0	0	3	3
Veneto	19	5	.26	3	2	3	0	0	3	3
Liguria	15	2	.13	1	0	0	0	0	2	4
Puglia	11	1	.09	0	0	0	0	0	1	1
Sicilia	10	2	.20	0	0	0	0	0	2	2
Friuli VG	9	3	.33	4	3	2	2	3	1	2
Marche	6	3	.50	3	2	1	3	3	3	3
Toscana	5	0	–	2	0	0	0	0	0	0
Calabria	4	0	–	0	0	0	0	0	0	0
Trentino AA	3	1	.33	1	1	1	2	2	0	0
Basilicata	3	1	.33	0	0	0	0	0	2	2
Abruzzo	6	0	–	0	0	0	0	0	0	0
Sardegna	2	0	–	1	0	0	0	0	0	0
Umbria	1	0	–	0	0	0	0	0	0	0
Molise	0	0	–	0	0	0	0	0	0	0
Val d'Aosta	0	0	–	0	0	0	0	0	0	0
Totals	292	83	.28	37	26	69	30	39	30	39

3. Methodological notes for the ego-network analysis

We conducted in-depth interviews with startuppers "with social vocation" to bring out the strategies they activate in their ecosystem and reconstruct from below the organizational field. The study is conducted according to a qualitative approach where the interviewee's words become *network narratives* which "provide an account of events and experiences and the ways in which they are connected together from an actor's point of view" (Crossley *et al.*, 2015, p. 106). At an operational level, we wish to check through their perceptions:

- whether they weave relationships with organizations of different social nature;
- how they position themselves with respect to organizations and institutions on the different territorial levels;
- whether they define their companies as immersed in a dense network.

In other words, we wish to check if there were conditions for a change in the welfare system from the *vertical* and *horizontal* to the *circular subsidiarity*.

We stimulated the startuppers to reconstruct their relationships with *relevant alters* such as supporters, trainers, partners in social impact activities, significant persons to whom they recognize the possession of specific skills concerning issues of social impact and social innovation, and/or a particular charisma. We used an *open-ended approach*, leaving them free to indicate how many names they prefer.

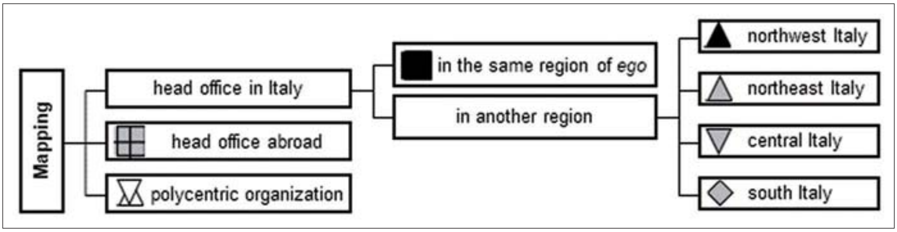
This approach "has the advantage of allowing the researcher an opportunity to derive a better sense of ego's network size and may also reach further into ego's network, beyond their immediate circle and towards weaker ties" (Crossley *et al.*, 2015, p. 51). The con is the difficulty in establishing boundaries which allow a better comparison between the different ego-networks. The depth of the analysis depends above all on the interviewee's willingness to tell their experiences. Some of them are elusive and respond briefly, others (the vast majority) are more talkative. The interviewer's skill is about being proficient in soliciting the former and stemming the latter, containing biases related to self-promotion and *social desirability*. The interview guide allowed us to distinguish, among the different types of relationships:

- the *strong ties* of *ego*, which are ties with those organizations or people without whom the startup would not have existed or, otherwise, would have difficulty remaining in business;
- the *problematic relationships*, due to either the negative judgment of *ego* against *alter* or to the interviewee's perception of any prejudices of *alter* towards *ego*.

We have chosen not to indicate the direction of ties because of the peculiar fluidity of the interactions between *egos* and *alters* that emerged by field research. Once a contact has been established, startuppers tend to enrich it with new content to build new opportunities. Furthermore, it could be that the entrepreneur's social vocation is associated with a preference for personal rather than institutional relationships. From the quote of an interview, it is clear that this also depends on the size of the company "compared to when I was a manager in a large company, now I can be closer to people, in a more concrete way". As with *ego-alter* ties, also with *alter-alter* ties, we are interested in detecting whether the interviewee represents, in his/her mental horizon, relationships in which he/she feels involved in pursuing social impact objectives. It may be that existing *alter-alter* relationships are not detected, either because they are considered irrelevant or because the interviewee is unaware of them[5]. However, we are not interested in the "objective" structure of the network. We want to know which place the startupper assigns to the SIaVS in the network constructed in the discursive process of the interview.

In the graphs, nodes are anonymous because we want to relate some attributes of the organizations and people involved in the network and not to reveal their identity. Anonymity also allows us to respect the interviewees' privacy, protecting the interest to keep their value judgments confidential. We featured nodes by geographical location and by type of organization or person. The taxonomy in Fig. 4 illustrates the operational definition of the property "head office of the organization". More detailed information, if required, are shown on the node label.

Fig. 4 – Geo-location of nodes



The other relevant characteristic of nodes is their attribution to one of the three *institutional spheres*, namely the *public authority* (PA), the *busi-*

[5] If we had adopted a name generator in an orthodox way, we would certainly have obtained denser networks, but at the cost of forcing the interviewee mind towards an excessive consideration of formal ties.

ness community (BC), and the *organized civil society* (CS). The conceptual structure that allows us to attribute all the nodes to one of these three spheres is itself a result of the research, which will be discussed elsewhere. In the following graphs, node labels correspond to specific concepts at the bottom line of the conceptual structure. The complexity of the attribution derives from the hybrid nature of many of the organizations in the networks. How to place an institution like Fondazione Con il Sud, a non-profit entity “born from the alliance between foundations of banking origin and the world of the third sector and volunteering”? We define banking foundations as non-commercial bodies in the BC, while volunteering is the heart of CS. We resolved this case by considering the civilian vocation of the institution as a priority, but the issue is still open. SIaVSs, social enterprises, and benefit corporation belong to BC, as commercial entities whose primary purpose is not the profit but the generation of positive social impacts. We have chosen these categories of analysis because it seems suitable for the current change, also due to the reform process that Italy has started in responding to the European challenge for measuring social impact. These changes are throwing into crisis consolidated categories of analysis, such as the notion of ONLUS (non-profit organization of social utility) and, more generally, the distinction between profit and non-profit businesses. In the graphs, we trace a rectangle around nodes to indicate their belonging to the sphere of PA, an ellipse for BC, and a dashed ellipse for CS.

4. Ego-network analysis

In this essay, we focus on a few interviews which seem interesting to us both for their ability to illustrate the differences between networks in Campania and Piemonte, and for the issues emerging from network narratives.

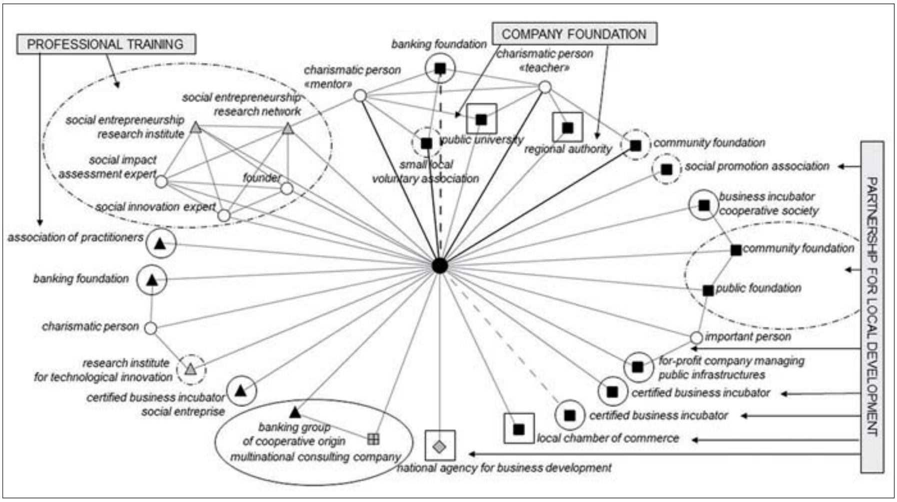
C01 was born on the basis of the experiences of a small voluntary association founded by the interviewee with other companions of a master’s degree at the university. The influence of two professors engaged in regional social policies (“mentor” and “teacher”) allowed the original idea to mature in a business project with a social vocation.

The startup was immediately acquired by a banking foundation and a community foundation, taking the first steps in a structure of ties consistent with the circular subsidiarity. C01 has never followed CIs acceleration paths, but it was at the center of a network involving incubators, accelerators and foundations from Campania and other regions. Despite the good results, the startup failed to be financially autonomous in short times. Activities were stopped abruptly due to internal problems of the banking foundation. The interviewee interprets the closure of the company more as a missed opportunity for the whole local system rather than as a personal failure.

Tab. 2 – General information about the selected cases and relative ego-networks

Cases	*Region*	*Interviewed*	*Sector*	*Size*	*Density*	*Components without ego*	*Homophily*
C01	Campania	GM	ICT	22 org + 7 persons	.13	4 + 7 isolates	BC 41% CS 41% PA 18%
C02	Campania	CEO; COO; social impact manager	ICT	25 org + 2 persons	.14	3 + 7 isolates	BC 24% CS 56% PA 20%
P01	Piemonte	CEO	Education	16 org + 2 persons	.21	2 + 5 isolates	BC 71% CS 24% PA 5%
P02	Piemonte	CEO; social media manager	Non-residential social assistance	23 org + 1 person	.20	2 + 2 isolates	BC 48% CS 38% PA 14%

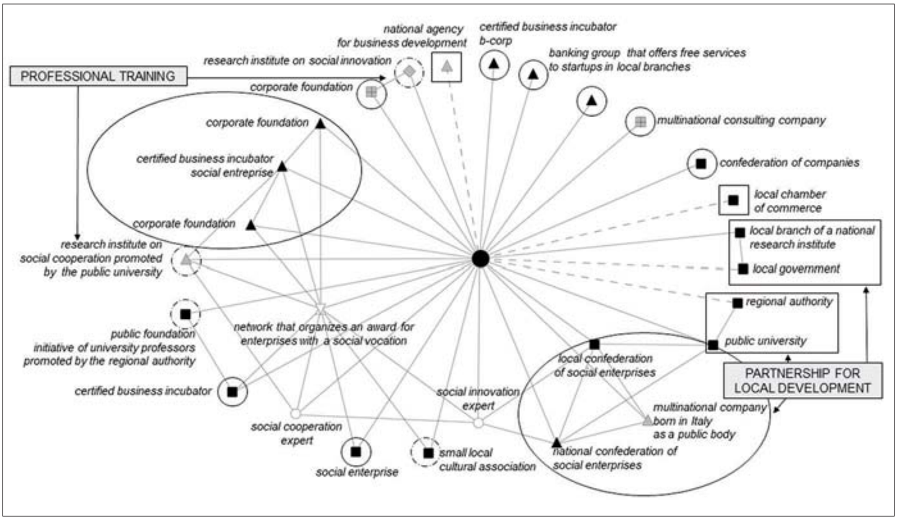
Fig. 5 – Network of C01



C02 is an *early-stage* startup run by men and women under 40. The start-up was born from an idea of the CEO, who called with him young professionals experienced in different fields. One has a strong background in the Third Sector as a professional and as a volunteer, one is a freelance expert in innovative technologies, one is a "nerd" engaged in civic participation, the

CEO comes from experiences in for-profit companies in the Northern Italy and returned to his homeland in Campania to realize himself and his social vocation. C2 startuppers report no “strong ties” with alters. To the question “is there anyone who supported you in starting the business?”, the CEO replied promptly: “Yes, *in primis* out of our own pocket. Then we won some national prizes and competitions dedicated to social innovation”.

Fig. 6 – Network of C02



P01 is a predominantly female startup founded by entrepreneurs with a highly professional profile, who have left fairly stable and remunerative professions to devote themselves to their social vocation. The interviewee tells us that she “heard the classic call to arms” after significant experiences in the family and in voluntary associations which pushed her to her previous job “and all the related comforts”. P01 has a strong tie with a multinational company that owns the patent. They have undertaken to spread their activities both by promoting the birth of new startups and sharing the project with voluntary associations all over Italy.

Like C01, P02 arises from the encounter between a small voluntary association and a local banking foundation which decides to invest in the ideas of some young people for the development of the territory. P02 is immediately following acceleration paths with a CI specialized in social impact projects. While having some success, they do not feel very comfortable in the world of business and finance. They come, in fact, from educational, professional, and voluntary paths that bind them in a rather tight way to the world of the Third

Sector and civic participation. The CEO immediately explains: "I am a man from the Third Sector. And I think our startup is still from the Third Sector. It is a third sector that questions how to find new challenges".

Fig. 7 – Network of P01

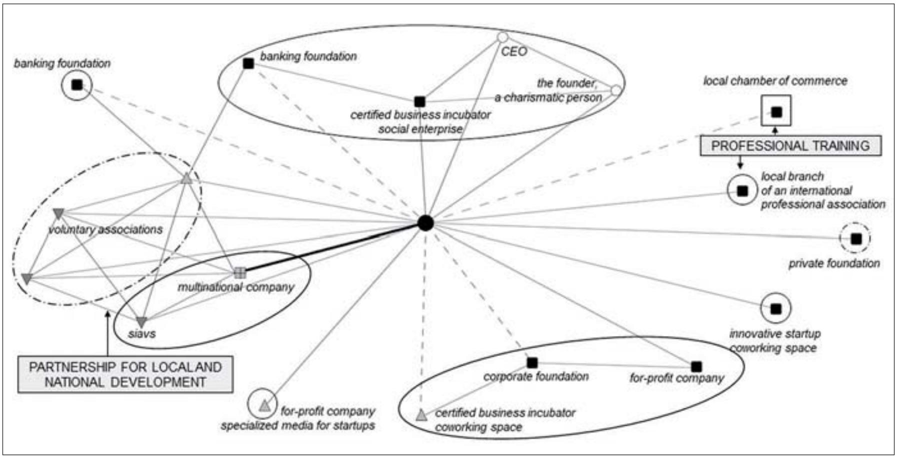
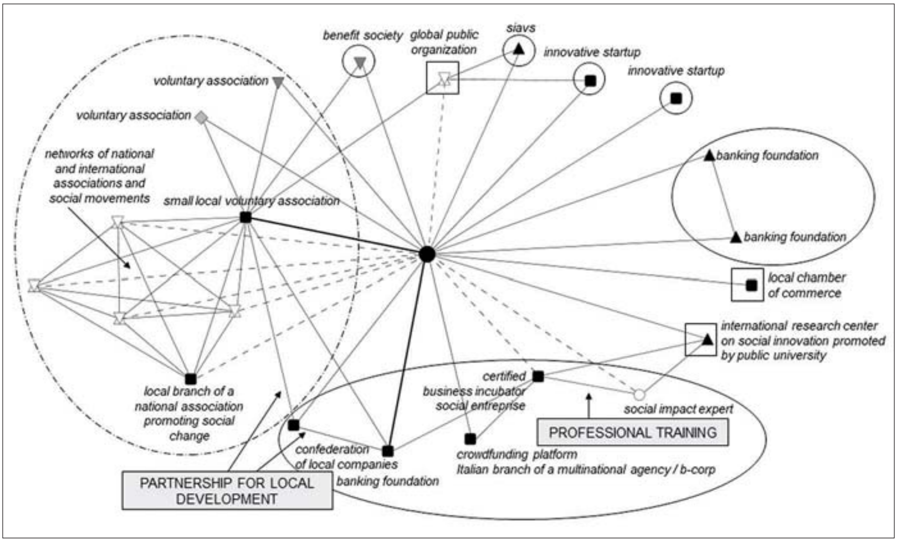


Fig. 8 – Network of P02



Despite the significant differences in individual stories, the structures of the two SIaVSs from Campania have elements in common; the same goes for the two from Piemonte. In Campania we find lower densities due to many one-to-one relationships. Dropped ego, the nets are divided into 3 or 4 components and 7 isolated nodes. In both cases, the larger component holds onto a *cut point* that is a single person, not an organization. For C01 he is the "mentor", a charismatic person with a great relational capital; for C02 he is a social innovation expert who is very active on the topics of social cooperation and social impact throughout Italy. The same person was nominated by both SIaVSs.

About the SIaVSs from Piemonte, their ego-nets have similar density but P02 has a larger size. Both show a clique where voluntary associations are involved. In the case of P02, they are large association networks and international movements in which wide-ranging public issues are debated, for P01, they are smaller associations that carry out the same project as the SIaVSs with voluntary activities. Both have more or less complicated relationships with the network of banking foundations and incubators. Two shared themes emerge strongly from the interviews and leave visible traces in the ego-networks. One is the *Southern question*, the other is the *hybrid nature* of SIaVSs which places startuppers in the complicated role of *mediator between PA, BC and CS*, each sphere with its own language and practices.

Southern startuppers perceive a profound inequality of opportunity between North and South. A first point is the lack of entities that make capital available to startups. C01 said: "there are 88 banking foundations in Italy, but they are almost all in the North". C02 said: "here there are no big companies and big foundations capable of supporting innovation and startups". We have seen that in the south there is also a lack of CIs that favor the meeting between startups and capitals (Figures 1-3). Both C01 and C02, which have no incubation ties with Northern CIs, have engaged relationships with some of them: "in Turin and Milan we met business accelerators specialized in social issues. There, we saw a whole network that based its development model precisely on the social impact" (C02).

Perceptions in Piemonte and Campania sound opposite. The CEO of P01 does not feel supported at all by foundations and accelerators. She says that her business partners in Campania have some advantages: "there are so many fundings dedicated to our business sector in Southern Italy. Here in the North there is an endemic lack of funds". Many of the *alters* from P01 network are members of "Torino Social Impact", but the CEO is unaware of it: "frankly, I don't know of any local networks dealing with social impact. Maybe the Turin Chamber of Commerce, but they are all networks where you must be the one who continually solicits to participate. Being behind these interlocutors is a heavy and expensive job. SIaVSs are left totally alone, or at least we have left us totally alone".

P02 knows this network and has a rather critical opinion on it: “In Turin, fascinating things are moving on social innovation. It’s okay if public authorities wonder about some social issues by dialoguing with other realities. However, they are not paths taken with the genuine will to be around a table. They are still not facilitating processes to discuss which direction to give to change by listening to everyone’s voice”. The criticism of the PA is common to all interviewees, who complain about the lack of long-term vision. “The theme is to understand how politics want to concretely support SIaVSs not only from an economic point of view, but also from a policy sensitive to certain issues. The social impact challenge lies precisely in creating opportunities in the territories” (C01). It is, therefore, not a mere lack of funds.

The problem is how to direct public and private resources to general interest objectives in the most appropriate way: “although many economic resources reward companies from the South (because everyone is concerned about southern Italy not growing), in the end, you are not supported in growth towards innovation and social impact” (C02). Southern interviewees blame the “cultural imprint of southern Italy”, primarily the permanence of personalities in dealing with the public administration: “often the interaction with the public sector is still mediated by relationships of direct acquaintance. This situation also hinders our interaction with social workers on the territory. In Milan, instead, the dialogue with the PA is promoted by foundations that help startups to grow, companies are increasingly open to the territory, and civic activism is experienced in a different and more proactive way. There is a continuous breath of interaction that is unfortunately lacking in Southern Italy. If we have to improve something, this is precisely networking” (C02). The criticism of C02 towards the PA is found in the graph by the dashed ties (Fig. 6).

There also are obstacles of a bureaucratic nature, which affect the whole national territory. The qualification of a company “with social vocation” “is not helpful until things change at bureaucratic level. If the municipality cannot work with SIaVSs on social problems, how do they act on a territory?” (C02); in fact, the qualification “still does not allow you to participate in many of the calls dedicated to social policies” (P02). “A public body is wrong to refuse to collaborate with a ‘for-profit’ company. A non-profit organization may act in a predatory way, while a ‘for-profit’ company may be impact-oriented” (C02). P01 often choose to work with voluntary associations, renouncing the possibility of economic feedback for their commitment: “we’re struggling just to break even. In fact, we act as a small ONLUS. We provide an innovative service of general interest, which the public service is not prepared to offer at this time. We have to do it because we don't want Italy to fall behind. Our startup was born with this vocation, with this spirit of bringing innovation and growth in society, but our commitment must combine with profit. I know that in Italy, in certain sectors, to make a profit appears blasphemy, but then there are thousands of operators who make profits dressed as ONLUS" (P01).

Dialogue with CS is also often complicated. Again, Campania startuppers consider it a local problem. "In Milan there is a good feeling, while with the Naples office of the same association collaborating becomes more difficult» (C02). The social impact manager of C02 assumes: "perhaps, it's because of a prejudice they have. I come from the Third Sector. Often, they believe that the business, 'the profit' is a 'dirty thing'. Maybe they worry we will contact them to exploit their knowledge. I think not many people here know the difference between a traditional 'for-profit' and an impact-oriented company like us". The cleavage between CS and BC is evident in Figure 8. P02 is in a clique with the association from which it was born and a dense "network of network" of national and international associations. They too report that when they participate in meetings with CS as SIaVS they encounter resistance and prejudice: "If you are SRL and you talk about social impact finance there is an extreme distrust. We perceive an unwillingness to dialogue with 'strange beings' like us. They say, 'It was already difficult to talk among us, now we also bring those who are different from us and who want to make money'. As if finding economic sustainability strategies meant 'making money'".

On the BC side, they believe something new is being born around the social impact investing: "it's winking, it's interesting […] something new is being created, but it is still another world from the social, too financialized. Money is powerful, pressures are strong and there is a high risk of leaving behind your social vision". Their hope is that the issue of social impact could be "a brake on certain dynamics, a tool of control and surveillance to not lose your values". For this purpose, circular relationships should be activated, involving the three institutional spheres and putting the person at the center, respecting the diversity of opinions and visions: "we believe that the processes of change, of collective learning, must be slow and participatory. I believe it. There is a beautiful world that is moving from below. In my opinion, social impact is a good engine, but processes take time. We have to network and find new languages". All interviewees declare that integrating different actors in a cooperative network is one of the main goals. "Our territory offers many opportunities, but the operators are all disconnected. Our startup already connects several operators and interacts with institutions. Networking could be a significant factor in economic and social development" (C02). Startuppers are well aware that to generate social impact, understood as a radical change in the community, their commitment is important but not sufficient: "a transversal change should be declined, starting from the institutions, precisely to re-evaluate the way we act in the social. A good idea can stimulate and make people participate. It's from below that change can happen, but the institutions must do their part. There are beautiful ideas and a lot of commitment on the part of small businesses, precisely because it is hard to build something here. It is a pity that there is no vision by the institutions, above all: I mean by the regional administrations, the municipalities, the foundations".

Comparing the structure of the four ego-networks shows us that those from Piemonte are denser. Dropper ego and alters who are not organizations, the nets from Campania do not have cliques made up of more than three nodes. P01 built around its project a clique of five organizations, both from CS and BC. They hope to make it grow over time, but they perceive that they do not have any support from the institutions. P02 is within dense CS networks thanks to the voluntary association from which it was born, and it is in impact finance circuits as a startup, but they perceive great difficulties in mediating between the two worlds, blaming the poor support from the PA. C01 had started in an ideal condition for *circular subsidiarity*, in a collaborative network between organizations of the three spheres, but problems internal to the main stakeholder led to early closure. The interviewee says: "now the network is frozen, but it would take little effort to get back together". C02 is a young and vital enterprise. They take part in calls and have won some, thus entering relationships with BS and CS throughout Italy. They feel tough to mediate between these networks and local institutions, therefore building their own development space, yet they are working toward this impact goal looking for new opportunities.

5. Conclusion

This paper is intended to offer two kinds of contributions to the recently boomed field of research of social impact-oriented entrepreneurship. The first is thematic. Following some interesting suggestions from the recent Italian literature (Moro, 2009, 2020; Zamagni, 2016; Zamagni *et al.*, 2015), we connected the analysis of the innovative welfare networks to the issue of the change in the guiding idea of subsidiarity toward the circular subsidiarity. We adopt these categories to study a particular (and quite neglected) kind of innovative enterprise, the SIaVS, also presenting a brief overview of their development status and their relations with the equally new certified incubators. The second contribution is methodological. We adopt the ego-centric approach to social network analysis in support of the interpretative sociology and the qualitative approach to social research. We drew our graphs as *network narratives* (Crossley *et al.*, 2015) starting from the images and representations that we have drawn from the hermeneutic interviews with the startuppers "with a social vocation", following a little structured guide. We are confident that the ego-centric approach may allow the social researcher to make visible the difference between the structure of social bonds based on circular subsidiarity and that of the "traditional" horizontal and vertical subsidiarity. Of course, there are improvements to be made. There is the need to include the temporal dimension, to better qualify the nature of the ties, to better define the boundaries of the networks. Further, it would be appropriate to broaden the inquiry to more SIaVSs (as we are doing) and to other kinds of social impact-oriented businesses (and non) to enable comparisons on a rigorous basis and with adequate techniques.

The mainly qualitative approach and the number of cases presented here do not allow us to make any generalization about the possibility of social change related to the development of innovative welfare networks. However, the research highlights some factors which may favor or hinder the orientation of startuppers toward the circularity of social bonds.

Interviews and graphs give us a general picture of cultural resistance to the desired changes both from the perspective of the Civil Economy and from the perspective of the innovative welfare network promoted at the European level. If it is true that, as it seems in the early observations, the nodes that populate the networks around the SIaVSs tend to form a sort of corporative structure of ties among organizations of the same type, we should admit the persistence of the delegation requirements to the detriment of *circular subsidiarity*[6]. Compounding the problem is the fact that the overall mechanism of access to resources does not overcome the rigid dualism profit/non-profit. Actually, interviewees complain about the difficulties spotting in the cognitive maps of their interlocutors an arrangement consistent with both the social and entrepreneurial vocation. At the level of methods and practices, this contradiction could lead to putting competition for resources ahead of the general interest.

Having focused the observation on SIaVSs in Piemonte and Campania has allowed us to draw elements for future evaluation of the southern question. The matter seems to settle on this distinction: in the North, there is a greater concentration of social impact investors and business incubators; in the South, there seems to be more attention on the part of public funds for Third Sector social activities. Having to choose either one or the other path due to the geographical position certainly does not favor either the circular subsidiarity

[6] The hypothesis of network homophily should be tested with specific techniques and an adequate sample.

network or the development of an innovative welfare network, or, by going to the bottom of the question, the achievement of economic and social cohesion.

References

Bishop M., Green M. (2008), *Philanthrocapitalism: How the Rich Can Save the World*, Bloomsbury Press, London.

Borgatti S.P., Everett M.G. (1997), "Network analysis of 2-mode data", *Social Networks*, 19, 3, pp. 243-269.

Calderini M. (*ed.*) (2019, *Tiresia Impact Outlook 2019*, Politecnico di Milano School of Management, Milano, http://www.tiresia.polimi.it/uncategorized/tiresia-impact-outlook-2019-il-report-della-ricerca/ consulted on 02/24/2020.

Castells M. (2000a), "Materials for an exploratory theory of the network society", *British Journal of Sociology*, 51, 1, pp. 5-24.

Castells M. (2000b), "The rise of the network society. Economy, Society and Culture", in *The Information Age: Economy, Society and Culture. Vol. 1*, Wiley Blackwell Publishing, Hoboken.

Crossley N., Bellotti E., Edwards G., Everett M.G., Koskinen J., Tranmer M. (2015), *Social network analysis for ego-nets*, Sage, London.

Delanty G., Rumford C. (2005), *Rethinking Europe. Social theory and implication of Europeanization*, Routledge, New York.

Giddens A. (2007), *Europe in the Global Age*, Wiley, Hoboken.

Jensen O.B., Richardson T. (2004), *Making European Space: Mobility, Power and Territorial Identity*, Routledge, London.

Mazzuccato M. (2013), *The Entrepreneurial State. Debunking Public vs. Private Sector Myths*, Anthem Press, London.

Moro G. (2009), *Citizens in Europe. Civic Activism and the Community Democratic experiment*, Springer, New York.

Moro G. (2020), *Cittadinanza*, Mondadori Università, Milano.

Prodi R. (2000), *Towards a European civil society*, Speech commencing second European Social Week, Bad Honnef, 6 April, https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_00_124 consulted on 02/20/2020.

Richardson T., Jensen O.B. (2000), "Discourses of Mobility and Polycentric Development", *European Planning Studies*, 8, 4, pp. 503-520.

Walzer M. (1995), *Toward a Global Civil Society*, Berghahn Books, New York.

Zamagni S. (2016), "L'evoluzione dell'idea di welfare: verso il welfare civile", *Quaderni di economia del lavoro*, 103, pp. 337-360.

Zamagni S. (2018), "Il Terzo Settore in transito. L'urgenza della normatività sociale", in P. Venturi, S. Rago, *Il Terzo Settore in transito. Normatività sociale ed economie coesive*, Aiccon, Forlì.

Zamagni S., Venturi P., Rago S. (2015), "Valutare l'impatto sociale. La questione della misurazione nelle imprese sociali", *Impresa Sociale*, 6, pp. 76-97.

11. *The Erasmus student mobility network*[*]

by Silvia Leoni[**], Luca De Benedictis[***]

1. Introduction

As many social phenomena, the student mobility of the Erasmus programme can be analysed through a network perspective.

The Erasmus programme, today Erasmus+, started in 1987, marking a fundamental moment for education and training in the European Union. Since its creation, more than 9 million people had participated in the Erasmus+ programme at its 30th anniversary in 2017 (European Commission 2015, 2018a). Created with the objectives of strengthening the European identity, increasing individual skills and, ultimately, people's employability, the programme, which allows for a period of study or internship in a foreign country, has become popular among university students who participate with an increasing number year after year. The programme also represents a cultural phenomenon, since for many European students it is the first life experience abroad and, therefore, an opportunity to familiarize with a different culture.

Given the success of Erasmus+, the European Commission has proposed to double the funds allocated to the programme for the 2021-2027 period in order to support the mobility of 12 million people, aiming to increase the participation of those students coming from a more disadvantaged background (European Commission, 2018b).

For its importance, the Erasmus programme has been studied in depth by the literature. In particular, two main research strands can be identified,

[*] An Italian version of this contribution has been published with the title "La rete di mobilità degli studenti Erasmus: topologia e caratteristiche", in G. Busilacchi, E. Cedrola (a cura di), *La forza delle reti*, Aracne, Rome, 2020.

[**] School of Business, University of Leicester, silvia.leoni@leicester.ac.uk.

[***] Department of Economics and Law, University of Macerata, luca.debenedictis@unimc.it; Department of Business and Management, LUISS, ldebenedictis@luiss.it.

respectively investigating i) the effect on competences and employability of participants in the programme (see for example Engel, 2010; Parey & Waldinger, 2010; Ballatore & Ferede, 2013; Bryla, 2015; Di Pietro, 2015; Ferri, 2019) and ii) the relationship between participation in the programme and the participants' feelings of multiculturalism and European identity (see for example Oborune, 2013; Jacobone & Moro, 2014). This work focuses exclusively on the Erasmus mobility for study purposes. Using data available from the European Union Open Data Portal[1], the Erasmus mobility for the academic year 2013/2014 is analysed using a network analysis approach. The aim is to explore the network of Erasmus students and study the relationships between the participating countries. The study of the directed and weighted network of Erasmus student flows allows identifying the most important nodes that host and send students abroad, providing an overview of the participation in the programme for the year analysed, and highlighting asymmetries in student participation.

The approach chosen is not entirely original. A study of the Erasmus mobility network can be found in Derzsi *et al.* (2011). The authors analyse Erasmus mobility in 2003, studying its directed and weighted network, where the nodes of the network represent European universities and links are the students flows between these universities. The study of the topological properties of the network shows that, contrary to what expected, the network does not show the characteristics of scale-free networks (see Barabási, 1999), but can be traced back to a small world type of network (see Watts and Strogatz, 1998) with the tail of its degree distribution well described by an exponential function. Breznik *et al.* (2020) also use a network analysis approach and provide a general overview of Erasmus mobility for study in the years 2007-2013, identifying three different groups of countries among the participants: good senders and receivers, good senders only and good receivers only. Restaino *et al.* (2020) expand the network analysis tools by adopting a blockmodeling approach on the network of countries involved in the programme both for study and internship. Böttcher *et al.* (2016) study the Erasmus mobility for the a.y. 2011/12 and identify a gender gap both by country and by study area. In particular, for almost all participating countries and several thematic areas, female students are over-represented. This gap in favour of women is also found in De Benedictis and Leoni (2020a, 2020b) which analyse the period 2008-2013. Their work shows that the gender bias in favour of female students is persistent along the years considered, however, a mild trend towards its

[1] https://data.europa.eu/euodp/en/home.

reduction emerges and suggests that this tendency could foster the convergence of male and female students flows.

The work has been structured as follows: Section 2 explains the functioning of the Erasmus programme, its objectives and its evolution over time; Section 3 provides a quantitative description of the Erasmus mobility for study purposes in the a.y. 2013/14; Section 4 describes the application of network analysis and presents its results; finally, Section 5 summarizes the main results of the analysis and offers some reflections on the structural characteristics of the Erasmus programme and possible future changes.

2. The Erasmus programme over time

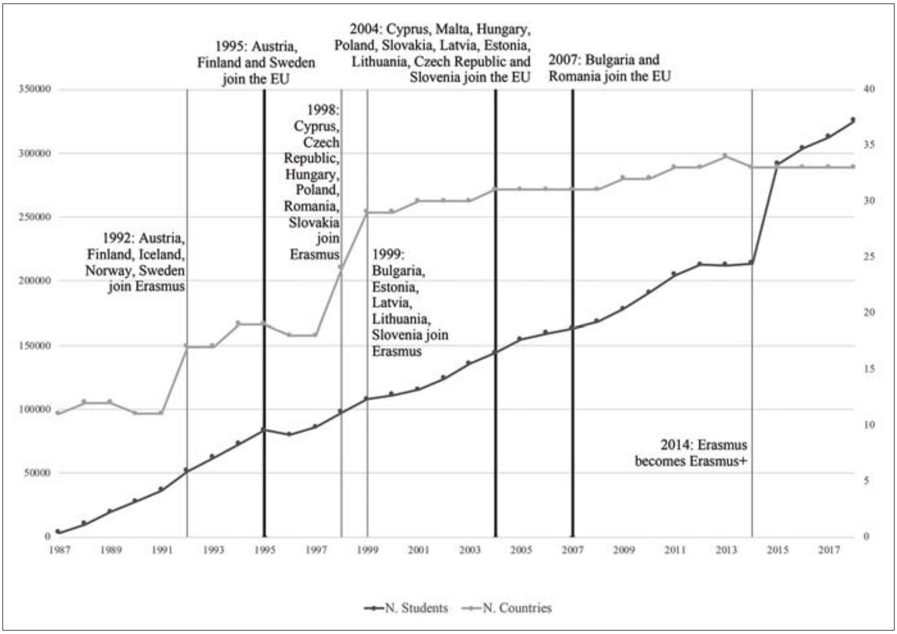
Erasmus is an acronym for the European Region Action Scheme for the Mobility of University Students. It is a student mobility programme of the European Union created in 1987. It started with the idea of offering to European students in higher education the possibility to study abroad for a limited period at a European university, legally recognized by their home university while being financially supported through a scholarship.

The evolution of participation in the Erasmus programme from 1987 to 2018 is depicted in Figure 1. In 1987, 11 countries and 3,244 students participated in the mobility programme for study purposes, while 2018 saw the participation of 33 countries and 325,495 university students. The growth in the number of participating students has followed the growth of countries' participation in the programme, which in many cases joined Erasmus before even becoming European Member States. This is the case of Austria, Finland, and Sweden, for example, which became EU Member States in 1995, but joined the Erasmus programme as early as 1992; or the case of countries that joined the EU in 2004 but have been participating in Erasmus since 1998 and 1999, such as Cyprus, Czech Republic, Lithuania, Latvia, Estonia, Slovakia. The number of students participating in the mobility saw an unprecedented increase in 2015, jumping from 213,879 to 291,383, although the number of participating countries remained almost unchanged. In 2014, the programme changed its name in Erasmus+ and acquired a more inclusive form (with the EU Regulation No 1288/2013) by allowing the participation of school and university teaching staff, and administrative staff, and expanding its area of interest not only to education but also to training, youth and sport.

In 2013 the programme reached the highest number of participants, 34 countries, made by the 28 EU countries plus Switzerland, Iceland, Liechtenstein, Norway, North Macedonia, and Turkey. Since 2014, Switzerland no longer enjoys the status of *member country* of the programme, but it is a *partner country*; i.e., it has adopted a transitional solution financed with Swiss funds that still allows Swiss people and institutions to participate in the program. In 2019, the programme countries were again 34 with the official inclusion of Serbia.

Thus, the programme continues to grow with the aim of strengthening and becoming more inclusive.

Fig. 1 – Evolution of the Erasmus programme from 1987 to 2018



Note: The darker line represents the number of students participating in the mobility programme (left scale), the light grey line represents the number of countries participating in university mobility (right scale). The thin vertical lines signal new countries joining the Erasmus programme and the transformation of the programme into Erasmus+. The thick vertical lines highlight some of the main enlargement of the European Union.

Note that the value for the year 2014 is not an ex-post figure, but a projection (European Commission, 2015).

3. A quantification of the Erasmus programme in 2013-14

The data used in this analysis are provided by the European Commission and available with free access at the EU Open Data Portal.

The most recent dataset refers to the academic year 2013-2014 and it is the one employed in this analysis. The dataset collects information on both study and placement mobility. This study focuses exclusively on mobility for study purposes, therefore observations related to internships have been excluded from the original dataset, resulting in a dataset containing 212,208 observations corresponding to the number of students who have benefited from the Erasmus grant. For the a.y. considered, the programme involved 34 countries, the 28 Member States plus Iceland, Liechtenstein, North Macedonia, Norway, Turkey, and Switzerland. The universities involved in the exchanges amounted to 2,732.

Table 1 provides some information on the participating countries, in particular the distribution of universities in the 34 countries, the distribution of incoming and outgoing students, and finally the percentage of female students who left and arrived for each country[2]. Moreover, two standardized measures for inbound and outbound flows are indicated in parentheses. In particular, for the incoming flows this measure is the percentage ratio between the number of incoming students and the average number of students enrolled in universities abroad; for the outgoing flows the measure is the percentage of outgoing students out of the total number of students enrolled in higher education in the country of reference.

For the incoming flows, this measure highlights the role of countries such as Spain, France, Great Britain, and Italy as destinations for European university students; for outgoing flows, on the other hand, the standardized measure shows the great incidence of Erasmus mobility for very small countries, such as Liechtenstein and Luxembourg.

Table 1 suggests that the countries involving the largest number of universities are those with the most numerous Erasmus student flows; this relationship is confirmed by the correlation coefficient between the variables. For both incoming and outgoing flows, the value of the correlation coefficient is close to 1 and therefore confirms that larger numbers of participating universities are associated with larger flows of students, especially outgoing (Pearson correlation coefficient is equal to $\rho_O = 0.88$) but also incoming (Pearson correlation coefficient is equal to $\rho_I = 0.83$).

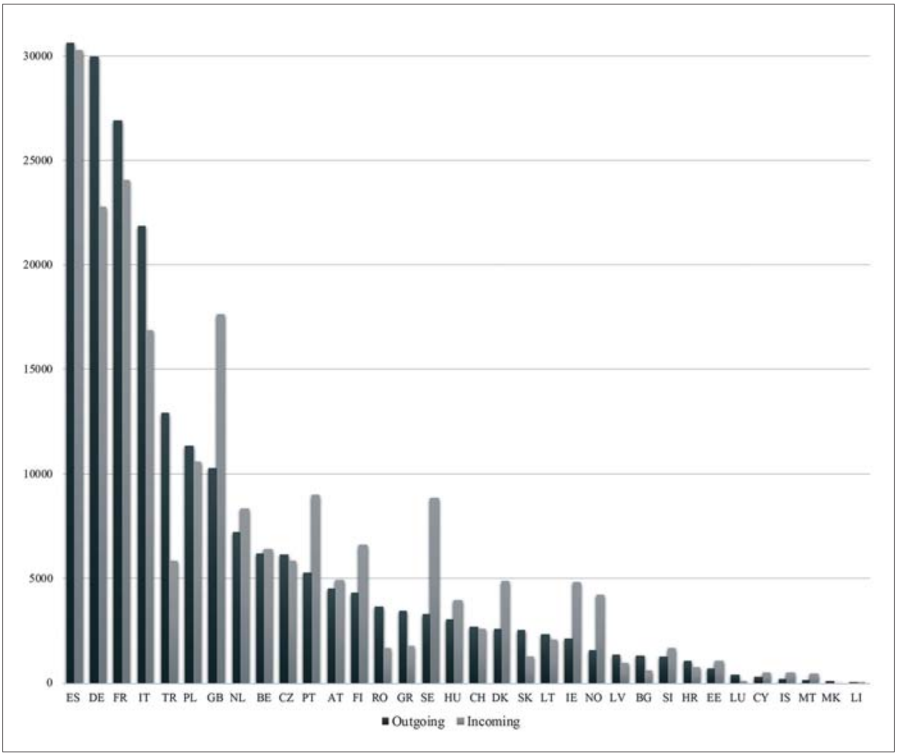
[2] In all tables and figures countries are named using the two-letter ISO code as follows: AT: Austria; BE: Belgium; BG: Bulgaria; CH: Switzerland; CY: Cyprus; CZ: Czech Republic; DE: Germany; DK: Denmark; EE: Estonia; ES: Spain; FI: Finland; FR: France; GB: Great Britain; GR: Greece; HR: Croatia; HU: Hungary; IE: Ireland; IS: Iceland; IT: Italy; LI: Liechtenstein; LT: Lithuania; LU: Luxembourg; LV: Latvia; MK: Macedonia; MT: Malta; NL: Netherlands; NO: Norway; PL: Poland; PT: Portugal; RO: Romania; SE: Sweden; SI: Slovenia; SK: Slovakia; TR: Turkey.

Tab. 1 – Distribution of universities and students' participation by country (2013-2014)

Country	*Universities*	*Incoming students*	*Outgoingstudents*	*F % incoming*	*F % outgoing*
AT	65	4,934 [0.66]	4,556 [1.08]	61.67	63.48
BE	76	6,402 [0.85]	6,247 [1.28]	63.53	62.64
BG	43	612 [0.08]	1,305 [0.46]	55.72	66.13
CH	34	2,611 [0.35]	2,702 [0.97]	53.77	58.59
CY	15	527 [0.07]	285 [0.89]	65.09	71.58
CZ	58	5,848 [0.78]	6,193 [1.45]	49.40	64.73
DE	321	22,809 [3.36]	29,983 [1.08]	58.89	61.09
DK	61	4,911 [0.65]	2,581 [0.89]	57.48	61.53
EE	20	1,059 [0.14]	716 [1.10]	55.43	71.23
ES	180	30,275 [4.30]	30,621 [1.55]	66.08	55.29
FI	48	6,618 [0.88]	4,339 [1.40]	54.49	63.01
FR	487	24,057 [3.47]	26,921 [1.15]	68.38	57.07
GB	169	17,654 [2.55]	10,282 [0.43]	64.60	64.01
GR	39	1,763 [0.24]	3,456 [0.52]	59.22	71.04
HR	24	746 [0.10]	1,070 [0.65]	58.18	67.85
HU	50	3,967 [0.53]	3,059 [0.85]	54.55	62.37
IE	34	4,821 [0.64]	2,121 [1.06]	62.52	60.73
IS	7	487 [0.06]	194 [1.02]	57.70	56.19
IT	182	16,872 [2.39]	21,889 [1.17]	65.01	58.71
LI	1	36 [0.005]	25 [2.96]	61.11	24
LT	42	2,082 [0.27]	2,327 [1.46]	46.01	67.86
LU	1	109 [0.01]	431 [6.51]	60.55	54.99
LV	38	976 [0.13]	1,367 [1.45]	47.44	70.52
MK	7	0	89 [0.15]	–	61.80
MT	3	480 [0.06]	151 [1.20]	66.04	58.94
NL	54	8,368 [1.13]	7,231 [1.07]	57.00	60.64
NO	49	4,226 [0.56]	1,558 [0.61]	57.48	61.62
PL	223	10,593 [1.50]	11,384 [0.60]	46.02	70.60
PT	89	9,020 [1.20]	5,325 [1.48]	60.45	52.66
RO	68	1,704 [0.23]	3,683 [0.60]	40.73	71.25
SE	42	8,874 [1.18]	3,324 [0.76]	51.10	60.11
SI	21	1,677 [0.22]	1,277 [1.31]	54.50	65.62
SK	33	1,274 [0.16]	2,568 [1.23]	50.92	66.20
TR	148	5,843 [0.95]	12,948 [0.26]	57.33	51.59

Note: For each of the 34 participating countries, the table shows the number of participating universities, the number of incoming students, the number of outgoing students, the percentage of incoming and outgoing female students (F) respectively on the total flows for the country of reference. In the column showing the number of incoming students, the percentage ratio between the number of incoming students and the average number of students enrolled in foreign universities is also shown in brackets. The column showing the number of outgoing students reports in parentheses the percentage of outgoing students out of the total number of students enrolled at the university in the country of reference.

Fig. 2 – Distribution of incoming and outgoing students by participating country, sorted in descending order of outgoing flows (2013-2014)



Countries that can be defined as *senders*, i.e. from which the greatest number of students leaves, are in order Spain, Germany, France, Italy, Turkey, and Poland. It is clear from Figure 2 that displays the distribution of outgoing students (dark grey bars). The light grey bars instead indicate the number of incoming students for each country. The first receiver countries only partially correspond to the sender countries. Spain is the country that sends more students on Erasmus mobility, but also receives more, followed by France, Germany, Great Britain, Italy, and Poland. It is worth noting the important role of Turkey as a sender, although it is not a member country of the EU, and the role of Great Britain as a receiver, which receives many more students with respect to the ones leaving from this country. Poland certainly stands out both in the role of sender and receiver among Eastern European countries.

Except for Liechtenstein, the majority of outgoing students for all participating countries are women, with percentages reaching around 71% for Cyprus, Estonia, Greece, and Romania. The same holds for the incoming flows, with the exception of a few Eastern countries, which mostly receive male students.

Using the International Standard Classification of Education (ISCED) developed by UNESCO to facilitate international comparison between education-related indicators, Table 2 groups mobility flows by gender according to the fields of study coded by ISCED-F 2013. There are four study sectors for which mobility is very high (> 10,000 students): (i) Arts and humanities, (ii) Social sciences, journalism and information, (iii) Business, administration, law and (iv) Engineering, with higher participation of women in Arts and humanities and men in Administration.

Tab. 2 – Distribution of Erasmus students by field of study and gender (2013-2014)

Field of study (ISCED-F 2013)	*F*	*M*	*F/M*	*Total*
Generic programmes and qualifications	180	82	2.20	262
Education	5,844	1,383	4.23	7,227
Arts and humanities	33,098	11,106	2.98	44,204
Social sciences, journalism and information	19,693	11,105	1.77	30,798
Business, administration and law	32,190	23,256	1.38	55,446
Natural sciences, mathematics and statistics	6,116	4,887	1.25	11,003
Information and Communication Technologies	1,286	3,629	0.35	4,915
Engineering, manufacturing and construction	14,232	20,808	0.68	35,040
Agriculture, forestry, fisheries and veterinary	1,860	1,319	1.41	3,179
Health and welfare	9,152	3,719	2.46	12,871
Services	3,128	2,373	1.32	5,501
Unknown	1,003	759	1.32	1,762
Total	127,782	84,426	1.51	212,208

Note: For each field of study classified according to ISCED-F 2013, the table shows the number of female students (F) and male students (M) in Erasmus mobility, as well as the total number of students. The F/M column contains the ratio between female students (F) and male students (M) for each field of study.

Table 2 also shows the ratio of female to male students (F/M) for each area of study. To be noted is the number of female students in mobility in Education which is more than four times the number of male colleagues, and in the Arts and humanities sector where the number of female Erasmus students is almost three times higher. Conversely, female students are about one third of male colleagues in the ICTs and represent the minority also in Engineering.

4. Network Analysis

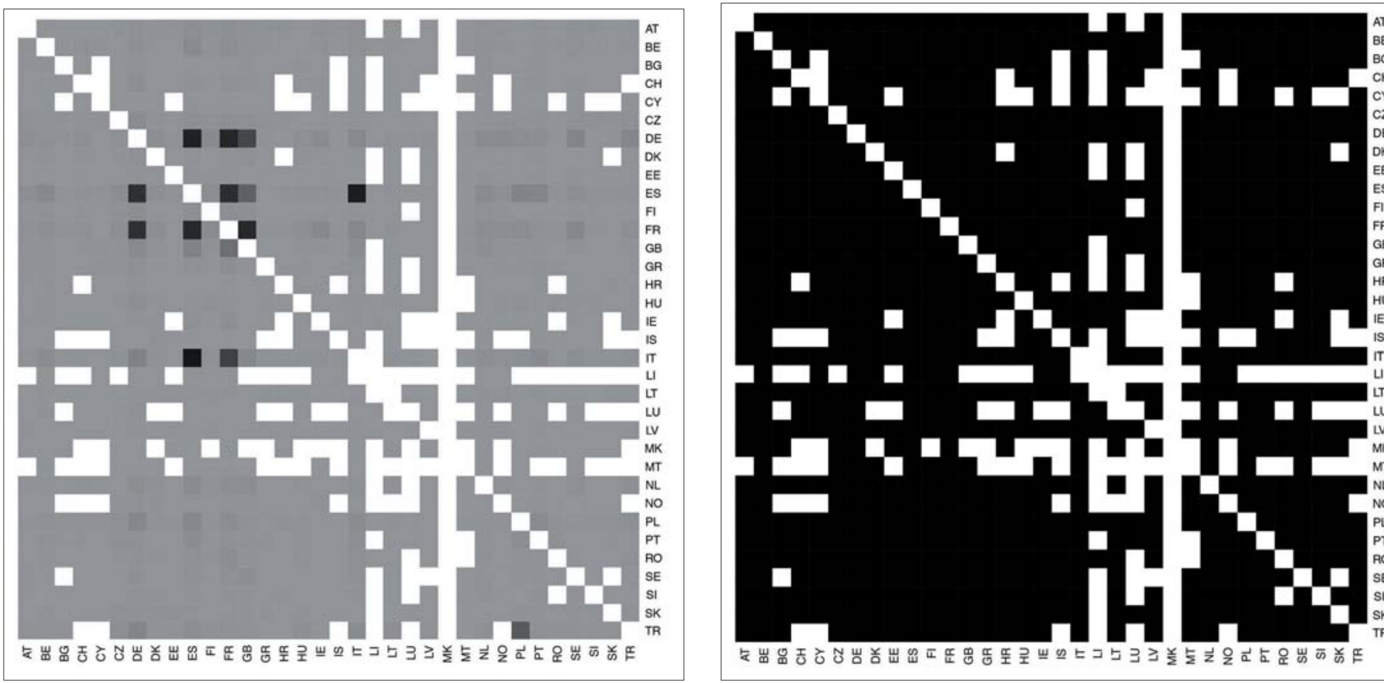
The elements highlighted by the previous analysis can be supplemented and reinforced by a simple application of *network analysis* to the flows of students participating in the Erasmus programme during the period considered. This is carried out in four phases. The first phase involves the definition of the reference network (e.g. the basic elements of the structure of students flows between countries). The second step is the extraction of the graph from the identified structure through the definition of an A_{ij}, adjacency matrix, where A is a binary matrix (i.e. composed of values equal to 1 or 0) that identifies the presence or absence of flows, while i defines the country of departure and j the country of destination of the student flows. The third one allows to visualize the network, projecting the adjacent matrix on the plane and using information on the value of the flows to generate the directed and weighted network of the Erasmus students, W_{ij}. Finally, the fourth phase consists in calculating some indicators that measure the relative position of the countries in the reticular structure of students flows and define their level of centrality.

The original data are those discussed in Section 3 and include the incoming and outgoing flows of students participating to Erasmus in the a.y. 2013-2014. This matrix, W_{ij}, has dimension n x n, where n is the number of countries participating in the Erasmus programme in 2013-2014, and includes w_{ij} elements that measure the flows of students leaving from country i and hosted in country j.

In Figure 3 countries are sorted in alphabetical order in both panels and identified by their respective ISO2 code. Figure 3a shows the *heatmap* of the W_{ij} matrix, while Figure 3b displays the A_{ij} matrix. Rows represent the countries of departure while columns the arrival countries. The values of the w_{ij} flows go from zero to the maximum value of 6,669 students departing from Italy to Spain. The colour scale in the visualization of the W_{ij} matrix goes from white (absence of flows) to dark grey, up to black as the values increase, while the A_{ij} matrix highlights the presence, in black, or absence of flows, in white.

By definition, w_{ij} flows along the main diagonal are zero, with no exchange between universities in the same country. The predominance of black cells over white cells in Figure 3b shows the degree of pervasiveness of the Erasmus programme's interconnectivity: many countries, like Belgium, for instance, or Czech Republic, send and receive students to and from almost all other countries, and this tends to be the predominant element of the Erasmus student flow structure. The white cells, on the other hand, highlight some specific cases.

Fig. 3 – W_{ij} matrix (left) and A_{ij} adjacency matrix (right) of the Erasmus student flows in 2013-14



(a) *(b)*

Note: The W_{ij} matrix (a) is represented in the form of heatmap where the grey scale takes darker intensity as the values increase, thus ranging from white, which indicates the absence of flows, to dark grey, up to black; the A_{ij} matrix (b) is represented in black and white and highlights the absence (white) or presence (black) of flows.

This evidence can be summarized by the value of the network density corresponding to the A_{ij} matrix. The density index, calculated as the ratio between the number of existing links (928, in our case) and the maximum number of possible links (n x (n-1) = 34 x 33 = 1,122) is 0.83. In other words, considering two random nodes in the Erasmus student network, the probability of these two countries to be linked is 83%.

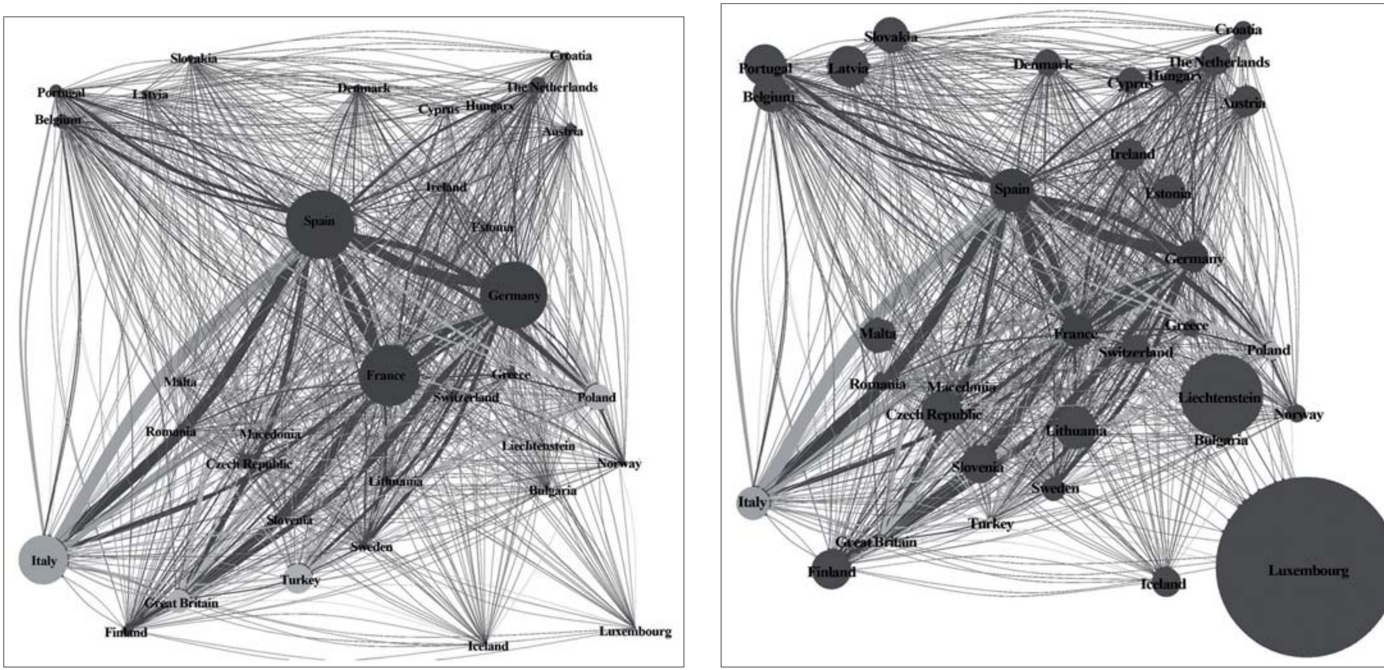
Contrary to the homogeneity in country participation both in terms of sender and receiver role, the visualization of the W_{ij} matrix shows that the degree of participation is very heterogeneous: the sum of the first 48 flows (out of 928 realized in 2013-2014) covers more than 50% of the number of Erasmus students; the sum of the first 302 flows covers 90% of it. The remaining 646 flows are therefore completely residual (with 204 bilateral flows with a w_{ij} value of less than 10 students).

This heterogeneity can be easily distinguished in Figure 4, representing the network structure. Participating countries are represented by the nodes of the network, while the links between them describe the Erasmus students mobility flows, directed in such a way as to distinguish the outgoing and incoming flows, and described by the arcs of variable thickness (proportional to w_{ij}) and of shades of grey similar to that of the country-node of departure.

The most connected countries in terms of incoming and outgoing links are displayed at the centre of the relations space by the Fruchterman-Reingold visualization algorithm (see Fruchterman and Reingold, 1991), while the nodes corresponding to the less connected countries are positioned in more peripheral locations. In Figure 4a, the nodes size is proportional to the number of outgoing Erasmus students.

The most relevant nodes in the network appear to be Spain, Germany and France. These three countries, with 30,621, 29,983 and 26,921 outgoing Erasmus students and 30,275, 22,809, and 24,057 incoming Erasmus students respectively, form the core of the network structure of Erasmus student flows. Italy alone, with 21,889 outgoing and 16,872 incoming Erasmus students, represents a second cluster, characterized by a strong bilateral connection with Spain, but less intense exchange flows with France and Germany. Finally, Poland, Turkey and Great Britain, although with a lower outflow of students (10,593 inbound and 11,384 outbound; 5,843 inbound and 12,948 outbound and 17,654 inbound and 10,282 outbound, respectively) constitute a third pole, but show a profound difference between them: Turkey shows a strongly positive outbound balance, while Great Britain, on the contrary, shows a negative balance. The fourth and last cluster, although marginal in terms of bilateral flows and showing very low overall values, is, on the other hand, the most numerous group of countries participating in the Erasmus programme.

Fig. 4 – Network of the Erasmus students flows in the a.y. 2013/2014



(a) *(b)*

Note: The nodes size is proportional to the number of outgoing Erasmus students (a) and to the percentage of university students participating in the programme (b). The thickness of links is proportional to the student flow for which the link is weighted. The grey shades indicate four clusters of different countries.

Tab. 3 – Centrality indexes of the country-nodes in the Erasmus student network in 2013-14

	Indegree	*Outdegree*	*Instrength*	*Outstrength*	*Relative instrength*	*Relative outstrength*
AT	0.91	0.88	0.16	0.15	0.15	0.17
BE	0.97	0.94	0.21	0.20	0.20	0.20
BG	0.76	0.82	0.02	0.04	0.02	0.07
CH	0.76	0.74	0.09	0.09	0.08	0.15
CY	0.76	0.56	0.02	0.01	0.02	0.14
CZ	0.94	0.94	0.19	0.20	0.18	0.22
DE	0.97	0.94	0.75	0.98	0.78	0.17
DK	0.91	0.82	0.16	0.08	0.15	0.14
EE	0.82	0.88	0.03	0.02	0.03	0.17
ES	0.97	0.94	1.00	1.00	1.00	0.24
FI	0.94	0.91	0.22	0.14	0.20	0.22
FR	0.97	0.94	0.79	0.88	0.81	0.18
GB	0.91	0.91	0.58	0.34	0.59	0.07
GR	0.82	0.88	0.06	0.11	0.06	0.08
HR	0.74	0.76	0.02	0.03	0.02	0.10
HU	0.85	0.91	0.13	0.10	0.12	0.13
IE	0.91	0.76	0.16	0.07	0.15	0.16
IS	0.71	0.56	0.02	0.01	0.01	0.16
IT	0.94	0.91	0.56	0.71	0.56	0.18
TR	0.76	0.76	0.19	0.42	0.22	0.04
LI	0.35	0.32	0.00	0.00	0.00	0.45
LT	0.85	0.91	0.07	0.08	0.06	0.22
LU	0.44	0.53	0.00	0.01	0.00	1.00
LV	0.76	0.94	0.03	0.04	0.03	0.22
MK	0.00	0.56	0.00	0.00	0.00	0.02
MT	0.68	0.38	0.02	0.00	0.01	0.18
NL	0.97	0.88	0.28	0.24	0.26	0.16
NO	0.76	0.71	0.14	0.05	0.13	0.09
PL	0.91	0.94	0.35	0.37	0.35	0.09
PT	0.91	0.88	0.30	0.17	0.28	0.23
RO	0.76	0.88	0.06	0.12	0.05	0.09
SE	0.94	0.79	0.29	0.11	0.27	0.12
SI	0.85	0.85	0.06	0.04	0.05	0.20
SK	0.74	0.91	0.04	0.08	0.04	0.19

Note: For each of the 34 countries participating in the Erasmus programme, the table shows three variations of the centrality index. Columns 2 and 3 indicate respectively the centrality index based on the sum of the incoming (indegree) and outgoing (outdegree) links; columns 4 and 5 indicate respectively the centrality index based on the connectivity weighted by the volume of the incoming (instrength) and outgoing (outstrength) flows; columns 6 and 7 show the values of instregth and outstrength in relation respectively to the total number of university students in the countries from which the reference country receives students and the total number of university students in the country.

The clusters can be distinguished by different intensity of grey in Figure 4. They have been identified through the maximization of *modularity*, which measures the number of edges falling within groups minus the expected number in an equivalent network with edges placed at random.

In Figure 4b, the nodes size is proportional to the percentage of university students in the reference country who benefited from the Erasmus programme. This different representation highlights the smaller countries, for which the participation in the Erasmus programme appears widespread among students in higher education.

In order to acquire information on the relative position of each country within the reticular structure, it is necessary to look at the local centrality indexes, commonly based on the *degree*, which is given by the sum of the links of a node; therefore, a node is all the more central the more it has relations with other nodes.

Table 3 summarizes three versions of the degree-based local centrality index. Column 2 and column 3 report the values of the centrality indexes based on the sum of the incoming (*indegree*) and outgoing (*outdegree*) links, both weighted with respect to the number of possible links for each node in the *n*etwork: n-1. For instance, being $n = 34$, Austria (AT), which receives students from 31 countries, has an *indegree* of 31/(34-1) = 0.91; Belgium, which receives students from 33 countries, has the highest value for the indegree.

In the case of *outdegree* and *indegree*, the countries with higher centrality do not systematically coincide, showing the presence of countries more active in welcoming foreign students and countries relatively more active in sending local students abroad. In addition to the countries mentioned above, the centrality measures highlight the central role of the Netherlands, Sweden, Finland and the Czech Republic in terms of hospitality. On the other hand, Slovakia, Poland, Latvia, Lithuania, Hungary, Finland and the Czech Republic are relatively central in terms of outgoing flows.

Columns 3 and 4 show the centrality indexes based on the degree weighted with respect to the volume of incoming (*instrength*) and outgoing (*outstrength*) student flows; more simply, the degree of each node is multiplied by the number of incoming or outgoing Erasmus students. The measures are then normalized by dividing by the maximum value of *instrength* and *outstrength*. These indexes highlight the predominant position of large countries such as Spain, Germany and France, Italy as well, but above all, the different condition of Turkey and Great Britain. The former has a value of *outstrength* twice the value of *instrength*, while the latter shows an opposite relationship between *outstrength* and *instrength*.

Finally, columns 5 and 6 relate the *outstrength* values with the total number of university students in the country and the *instrength* values with the total number of university students in the countries from which the host country receives Erasmus students. These are the same measures indicated in brackets in Table 2, again normalized to the maximum value. As noted in the lower panel of Figure 4, countries that manage to include a relatively high number of higher education students within the Erasmus programme are the very small ones, such as Lichtenstein and Luxembourg.

5. Conclusions

With its 33 years of experience, the Erasmus programme has become a growing European cultural phenomenon, which has proven to have a positive impact on the employability of European students and has become a substantial requirement for entry into the labour market.

Being a system of international relations with a reticular structure, it is a phenomenon that can easily be analysed with a network analysis approach, where the participating countries represent the nodes of the network and the flows of students are the links between countries. Through a graphical analysis and some network statistics, this work analyses the phenomenon of Erasmus for the academic year 2013/2014, providing a more recent evidence with respect to the research of Derzsi *et al.* (2011) and Böttcher et al. (2016) which respectively analyse the Erasmus students mobility network in 2003 and 2011. Unlike Restaino *et al.* (2020), who also study the mobility for internships, the network studied in this work only considers mobility for study reason.

The chosen methodology allows to highlight possible asymmetries in the network and the main countries in terms of incoming and outgoing flows, similarly to the research of Breznik *et al.* (2020). The study reveals some heterogeneity. In terms of flows, the network analysis allows to identify four clusters of countries with different degrees of participation in the programme. Among these, the cluster of countries with the highest level of incoming and outgoing flows is represented by Spain, Germany and France. Italy, characterized by a strong bilateral connection with Spain, represents the second cluster for the number of flows. The third cluster includes Poland, Turkey and Great Britain, among which Great Britain plays mainly a receiver role, while Turkey plays a sender role.

Looking at the flows of students in relation to those enrolled in tertiary education, the impact of the Erasmus programme is striking in small countries such as Liechtenstein and Luxembourg, which have very high shares of outgoing students with respect to their university students.

The centrality measures confirm the predominant position of the countries identified in the clusters through the graphic analysis, but also highlight the active role of other countries in terms of hosting students from abroad, as in the case of the Netherlands, Sweden, Finland and the Czech Republic, and in terms of outflows, as for Slovakia, Poland, Latvia, Lithuania, Hungary, Finland, and the Czech Republic.

Asymmetries are also observed in terms of disciplines of study, with a preference for Art and humanities, Social sciences, Engineering and Administration, which count more student flows than sectors such as ICTs or Natural sciences. Further differences can be found in terms of gender, with stronger female participation in the mobility.

The growth of Erasmus in terms of student participation suggests that the programme will grow further, also considering the increasing EU funding devoted to the initiative. On the other hand, the number of participating countries is rather stable but reflects the political dynamics of EU member and partner countries. For example, the recent exit of Great Britain from the European Union raises doubts about possible consequences for the Erasmus programme, which necessarily requires stability of agreements between universities to allow study abroad without incurring in the payment of tuition fees in the host country. Future research could explore possible effects that are still uncertain today.

References

Ballatore M., Ferede M.K. (2013), "The Erasmus Programme in France, Italy and the United Kingdom: student mobility as a signal of distinction and privilege", *European Educational Research Journal*, 12 (4), pp. 525-533.

Barabási A.L., Albert R. (1999), "Emergence of scaling in random networks", *Science*, 286 (5439), pp. 509-512.

Böttcher L., Araújo N.A., Nagler J., Mendes J.F., Helbing D., Herrmann H.J. (2016), "Gender gap in the ERASMUS mobility programme", *PLoS One*, 11 (2), pp. 1-8.

Breznik K., Skrbinjek V. (2020), "Erasmus student mobility flows", *European Journal of Education*, 55 (1), pp. 105-117.

Bryła P. (2015), "The impact of international student mobility on subsequent employment and professional career: A large-scale survey among Polish former Erasmus students", *Procedia-Social and Behavioral Sciences*, 176, pp. 633-641.

De Benedictis L., Leoni S. (2020a), "Gender bias in the Erasmus students network", March 20, *arXiv:2003.09167*.

De Benedictis L., Leoni S. (2020b), "Gender bias in the Erasmus network of universities", *Applied Network Science*, 5 (1), pp. 1-25.
Derzsi A., Derzsy N., Káptalan E., Néda Z. (2011), "Topology of the Erasmus student mobility network", *Physica A: Statistical Mechanics and its Applications*, 390 (13), pp. 2601-2610.
Di Pietro G. (2015), "Do study abroad programmes enhance the employability of graduates?", *Education Finance and Policy*, 10 (2), pp. 223-243.
Engel C. (2010), "The impact of Erasmus mobility on the professional career: Empirical results of international studies on temporary student and teaching staff mobility", *Belgeo. Revue belge de géographie*, (4), pp. 351-363.
European Commission (2015), *Erasmus+ Programme: Annual Report 2014*, European Commission, Brussels.
European Commission (2018a), *Erasmus+. Annual Report 2017*, European Commission, Brussels.
European Commission (2018b), *Investing in people. Making Erasmus even better*, Publications Office of the European Union, Luxembourg.
Ferri V. (2019), "Mobilità internazionale e condizione occupazionale dei laureati italiani", *Sinappsi*, 9 (1-2), pp. 71-89.
Fruchterman T.M., Reingold E.M. (1991), "Graph drawing by force-directed placement", *Software: Practice and Experience*, 21 (11), pp. 1129-1164.
Jacobone V., Moro G. (2015), "Evaluating the impact of the Erasmus programmeme: Skills and European identity", *Assessment & Evaluation in Higher Education*, 40 (2), pp. 309-328.
Oborune K. (2013), "Becoming more European after Erasmus? The impact of the Erasmus programme on political and cultural identity", *Epiphany*, 6 (1), pp. 182-202.
Parey M., Waldinger F. (2010), "Studying abroad and the effect on international labour market mobility: Evidence from the introduction of ERASMUS", *The Economic Journal*, 121 (551), pp. 194-222.
Restaino M., Vitale M.P., Primerano I. (2020), "Analysing International Student Mobility Flows in Higher Education: A Comparative Study on European Countries", *Social Indicators Research*, 149, pp. 947-965.
Watts D.J., Strogatz S.H. (1998), "Collective dynamics of 'small-world' networks", *Nature*, 393 (6684), p. 440.

 ISBN 9788835124603

COMPUTATIONAL SOCIAL SCIENCE

The Social Network Analysis perspective has proven the ability to develop a significant breadth of theoretical and methodological issues witnessed by the contribution of an increasing number of scholars and the multiplication of empirical applications in a wide range of scientific fields. One of the disciplinary areas in which this development has occurred, among others, is certainly that of computational social science, by virtue of the developing field of online social networks and the leading role of information technologies in the production of scientific knowledge. The complex nature of social phenomena enforced the usefulness of the network perspective as a wealth of theoretical and methodological tools capable of penetrating within the dimensions of that complexity.

The book hosts eleven contributions that within a sound theoretical ground, present different examples of speculative and applicative areas where the Social Network Analysis can contribute to explore, interpret and predict social interaction between actors. Some of the contributions were presented at the ARS'19 Conference held in Vietri sul Mare (Salerno, Italy) in October, 29-31 2019; it was the seventh of a biennial meetings series started in 2007 with the aim to promote relevant results and the most recent methodological developments in Social Network Analysis.

Giuseppe Giordano is associate professor of Statistics at the Department of Political and Social Studies of the University of Salerno (Italy).

Marialuisa Restaino is a lecturer in Statistics at the Department of Economics and Statistics of the University of Salerno (Italy).

Andrea Salvini is full professor of Sociology at the Department of Political Science of the University of Pisa (Italy).

FrancoAngeli
La passione per le conoscenze