Social Networks xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Connecting the dots: implementing and evaluating a network intervention to foster scientific collaboration and productivity

Valerio Leone Sciabolazza[a,1,*], Raffaele Vacca[b,1], Christopher McCarty[c]

[a] Department of Business and Economics, University of Naples, Parthenope, Italy
[b] Department of Sociology and Criminology & Law, University of Florida, USA
[c] College of Liberal Arts and Sciences, University of Florida, United States

ARTICLE INFO

Keywords:
Team science
Scientific collaboration
Network intervention
Rubin causal model
Community detection

ABSTRACT

This article presents the design and implementation of a network intervention to foster scientific collaboration at a research university, and describes an experimental framework for rigorous evaluation of the intervention's impact. Based on social network analysis of publication and grant data, an innovative type of research funding program was developed as a form of alteration of the university's collaboration network. The intervention consisted in identifying research communities in the network and creating a new collaborative relation between pairs of unconnected researchers in selected communities. The new collaboration was created to maximally increase the overall cohesion of the target research community. In order to evaluate the impact of the program, we designed a randomized experiment with treatment and control communities based on the Rubin Causal Model approach. The paper describes the intervention design, reports findings from the program implementation, and discusses the statistical framework for future evaluation of the intervention.

Introduction

Scientific production is increasingly the result of collaboration. A growing body of evidence shows that collaborative projects and scientific teams tend to dominate the academic research landscape across all branches of science, with lone authors being increasingly rare (Greene, 2007; Leahey, 2016; Wuchty et al., 2007). *Interdisciplinary* collaboration, in particular, is the subject of growing attention and investments by university administrations, governmental funding agencies, and research policy makers (Jacobs & Frickel, 2009). In recent years, the drivers, dynamics and outcomes of scientific collaboration have even become the central subject of new fields of research, such as the Science of Team Science (SciTS) (Falk-Krzesinski et al., 2011) and the Science of Science (SciSci) (Fortunato et al., 2018). At the same time, university administrations and research funders have explored a variety of programs and policies to stimulate interdisciplinary collaboration. These include funding initiatives targeted to interdisciplinary projects, such as the INSPIRE program of the US National Science Foundation (NSF) (National Science Foundation 2015) and the Interdisciplinary Research Consortia program of the US National Institutes of Health (NIH) (National Institute of Health, 2007); interdisciplinary training programs such as the NSF Integrative Graduate Education and Research Traineeship (IGERT); and interdisciplinary university fellowship programs (Sá, 2008).

This article aims to contribute to research on the design, implementation, and evaluation of programs and policies to promote interdisciplinary collaboration. We present a novel approach to the problem of facilitating team science, adopting an experimental framework based on the notion of social network intervention (Valente, 2012). Using social network analysis of publication and grant data at the University of Florida (UF), we designed and implemented an innovative type of pilot funding program conceived as a form of intervention on the University's collaboration network. The intervention consisted in identifying research communities in the University's collaboration network, and maximally increasing the cohesion of selected (treatment) communities by adding collaborative links with specific structural properties in their subnetwork. This article describes the intervention design, reports findings from the program implementation, and presents an analytical framework for rigorous future evaluation of the intervention.

The contribution of this paper is twofold. First, we present a network intervention program designed to overcome the constraints and path-dependency observed in the natural evolution of scientific collaboration networks. Research on collaborative behavior in science

* Corresponding author.
E-mail addresses: valerio.leonesciabolazza@uniparthenope.it (V. Leone Sciabolazza), r.vacca@ufl.edu (R. Vacca), ufchris@ufl.edu (C. McCarty).
[1] The first two authors are equal contributors.

https://doi.org/10.1016/j.socnet.2019.11.003

suggests that, when presented with incentives for collaboration, scientists tend to form teams within their professional comfort zone: they replicate past collaborations (Dahlander & McFarland, 2013) or select collaborators with largely overlapping background and expertise (Leahey, 2016). One reason for this behavior is that, although the documented benefits of interdisciplinarity are many (Larivière et al., 2015; Lee & Bozeman, 2005; Rawlings & McFarland, 2011; Shi et al., 2009; Wuchty et al., 2007; Yegros-Yegros et al., 2015), collaboration can be costly for individuals, especially when it crosses disciplinary boundaries: collaboration entails costs of communication and coordination, the need for translation between different styles of thought and research methods, and a number of administrative tasks associated with team management (Bikard et al., 2015; Cummings & Kiesler, 2005, 2007; Hackett, 2005). To reduce these costs, scientists often select collaborators in their same domain of knowledge, producing research that reinforces and consolidates existing areas of research (Leahey, 2016). Analyzing data on pilot grants awarded by the University of Florida Clinical and Translational Science Institute (CTSI), an NIH-funded interdisciplinary research institute in medicine, Vacca et al. (2015) showed that most CTSI awards produced teams confined within close departments in the University's Health Sciences Center. In the same study, a survey among UF faculty about their interest in potential collaborations with specific colleagues found that views were strongly dependent on the geodesic distance between potential collaborators in the University's collaboration network: the percentage of respondents who considered the collaboration as potentially successful dropped dramatically when the proposed collaborator was at larger distances in the network. In other words, researchers felt that collaborations outside their scientific comfort zone were more risky and likely to fail.

Second, we propose an experimental approach to accurately evaluate the impact of research collaboration on scientific productivity. This is a particularly significant issue given the ongoing debate on the relationship between collaboration and productivity in science (Chung et al., 2009; Hollis, 2001; Medoff, 2003). Although a large body of literature has found a positive correlation between the number of collaborations and the productivity of individual researchers (Cainelli et al., 2015; Colussi, 2017; Ductor, 2015), other studies have shown a negative impact of co-authored papers on productivity (Acedo et al., 2006; Hamermesh and Oster, 2002, Leahey et al., 2017, Lee et al., 2005, McCarty et al. 2013). These inconsistencies might be due to the use of observational data, while experimental studies enable more robust statistical analysis and rigorous evaluation. The strategy proposed in this paper for experimental evaluation of a network intervention contributes to the growing research on network experiments for the study of social behavior (for a review see Aral, 2016; Eckles et al., 2017; Walker & Muchnik, 2014). In particular, our results suggest an innovative procedure for the identification of meaningful treatment and control groups in network experiments.

The rest of the article is organized as follows. Section 2 reviews the notion of network intervention and provides an overview of our intervention and evaluation program. Section 3 presents the population and data considered, and details the intervention's steps and measures. Section 4 describes the experimental design and statistical framework proposed for the intervention's evaluation. Section 5 reports results from the program implementation, and the last two sections discuss the results and conclude the article.

A network intervention for team science

Notions and studies of social network interventions have become increasingly popular in the health and social sciences over the past twenty years (Valente 2012, Latkin & Knowlton 2015, Valente 2017). A network intervention is any program that uses network data and analysis to promote behavioral change in a population or to enhance individual or collective performance in an organization. Supported by research in social psychology, sociology, and behavioral sciences, the fundamental intuition behind these programs is that modifications in social environments tend to have a greater impact on individual behavior than interventions addressing individual characteristics and conditions alone (Cialdini & Trost 1998).

Valente (2012) provides an extensive literature review and a useful typology of network interventions, distinguishing four intervention types: (1) Selection of *individuals* with specific properties or positions in a social network, such as opinion leaders, to be recruited as intervention champions; (2) *Segmentation* programs, in which groups of people, often identified with network methods, are the intervention target or the channel through which the intervention is administered; (3) *Induction* programs, which stimulate interaction among already connected people to accelerate the diffusion of specific information or behaviors in a network; (4) *Alteration* programs, which change the existing social network by adding or removing nodes or links.

The most straightforward and popular strategy of network intervention consists in the selection of *individuals* in important network positions, typically based on centrality measures, and their recruitment as opinion leaders and change agents. Studies of this type of programs abound in the literature (see reviews by Valente & Pumpuang 2007 and Latkin & Knowlton 2015). In a recent example, Paluck et al. (2016) describe a "social influence" intervention to reduce student conflict in public middle schools. They show that the intervention's impact, as measured by indices such as numbers of disciplinary reports of student conflict, was significantly greater when the program champions were recruited among 'social referents', namely highly indegree central students in the school contact network (network of students 'spending time with' each other).

Segmentation is the most effective strategy in contexts in which people make decisions in groups rather than individually, are committed to adhere to group norms, and try to maintain group identity. In this case, the adoption of a new behavior results from group interactions and decisions. Buller et al. (1999) describe a network intervention that combined individual and segmentation strategies with the goal of increasing fruit and vegetable intake among blue-collar employees from ten public employers in Arizona. The program consisted in detecting cohesive subgroups (segmentation) and recruiting opinion leaders within subgroups (selection of individuals). Cohesive subgroups were identified as 2-cliques of employees in workplace contact networks (networks of co-workers "talking to" each other). Cliques were matched and randomized to treatment and control groups, and opinion leaders were recruited in treatment cliques to serve as peer educators and discuss eating fruits and vegetables with co-workers. Employees in treatment cliques showed significantly higher levels of fruit and vegetable consumption at the end of the intervention and in a follow-up survey six months later. In a design similar to Buller and colleagues' study, we identified cohesive communities in collaboration networks, matched and randomized communities between treatment and control groups, and administered the intervention through selected *dyads* (rather than individuals) in the treatment communities.

Our program combines elements of the *segmentation* and the *alteration* strategies: first, we use network science methods to identify cohesive groups or communities to be targeted by the intervention (segmentation); then, we facilitate new collaborations in selected dyads by adding edges in the identified communities (alteration). A similar combination of group-based intervention and creation of new ties is proposed by Gesell et al. (2013, 2016) in a program aimed at preventing childhood obesity. In this study, the intervention was administered to groups of 8-10 parent-child pairs (segmentation), and one of its components consisted in facilitating new social ties among participants (alteration). The aim was to produce a sense of individual integration and group cohesion, which were hypothesized to encourage the adoption of new healthy behaviors. The intervention was found to increase the number of nominated advice and discussion partners (outdegree) and network density, leading to an improved feeling of group cohesion (sense of belonging and group morale) in individual participants.

The setting for our intervention was the University of Florida, a US top-ten public research university that hosts a Health Science Center and sixteen colleges across the natural, health, and social sciences and the humanities. Designed as an experiment to be evaluated using the Rubin Causal Model (Rubin 1973, 1974), the intervention was implemented by means of an intramural pilot funding program supported by the UF Clinical and Translational Science Institute. Thus, the intervention can be regarded as the exploration of a novel approach to the design of pilot funding programs and the assessment of their impact. The intervention program took place in 2016-2017, and comprised the following steps:

- Detection of research communities. Publication and grant data from the previous three years (2013-2015) were analyzed to detect research communities at the University of Florida, as detailed by Leone Sciabolazza et al. (2017). Specific research communities were selected as intervention targets.
- Selection of potential collaborators within research communities. Based on network structural criteria, two potential collaborators were identified in each target community selected in Step 1. The goal of the intervention was to generate a collaboration between them in order to increase overall cohesion in the respective research community.
- Creation of the missing link between potential collaborators. In each community, the missing link between the potential collaborators selected in Step 2 was created through a UF CTSI limited submission pilot funding program.

To allow for statistical evaluation of the intervention impact, the program was designed as a randomized experiment with treatment and control units (**section 4**). Among the research communities detected in Step 1, treatment communities were selected and matched with control communities based on specific criteria. This article describes the design of the intervention and reports findings from Steps 1-3. Future research will be dedicated to analyzing the impact of the program using a combination of qualitative methods and quantitative analysis in the Rubin Causal Model framework.

Data and intervention

The population considered for this experiment is composed of the set of all the research communities detected by Leone Sciabolazza et al. (2017) in the publication and grant collaboration network of the University of Florida, and including at least one researcher in the health sciences. Intuitively, a research community is a (sub-) network of researchers that have consistently worked together from 2013 to 2015, the three-year period before the intervention took place.

Step 1: detection of research communities

We detect research communities analyzing data on peer-reviewed publications and externally awarded grants involving at least one University of Florida researcher in 2013-2015. Peer-reviewed publications are typically regarded as the main product of scientific activity and collaboration in academia (e.g., Long & Fox 1995, Lee & Bozeman 2005, Wuchty et al. 2007). Research grants represent an important gateway for academic science production, scientific collaboration, and future publications (Jacob & Lefgren 2011, Lee & Bozeman 2005), and are often viewed as an index of scientific productivity in addition to publications (Rawlings & McFarland 2011). This study combines publication and grant data to operationalize a broad, comprehensive notion of scientific collaboration in university settings, which is relevant to both types of research product. Our goal is to capture and intervene in realistic research communities as revealed by collaboration on both publications and grants. Publication data were obtained from the Web of Science database through VIVO, a semantic-web application adopted across all UF colleges and departments, which disambiguates author names and stores author and publication information (including unique numeric IDs for authors) for all UF researchers (Borner et al. 2012). Grant data were extracted from institutional records on financial transactions between funding agencies and the UF Office of Research. In these records, UF researcher names are already disambiguated and associated to the same numeric IDs used in the VIVO publication database.

Table 1
Structural characteristics of collaboration networks.

	G_{2013}	G_{2014}	G_{2015}
Number of nodes	4414	4038	3358
Number of edges	11950	9783	8363
Density	0.0012	0.0012	0.0015
Number of components	237	264	260
% of nodes in the giant component	85.09	81.08	77.43
Modularity of Louvain partition	0.8411	0.8558	0.8466
Number of Louvain clusters	280	306	298
Average size (SD) of Louvain clusters	15.76 (43.16)	13.19 (35.13)	11.26 (28.84)

We define a collaborative tie between researchers i and j as existing in a given year $t \in \{2013, 2014, 2015\}$ if: (1) i and j co-authored one or more articles published in that year; or (2) i and j co-appear as Principal Investigators (PIs), multiple PIs or co-PIs in the same grant(s) in that year. Thus, for each year t, we obtain a network $G_t(V, E)$ in which V is the set of all UF researchers who have collaborated on publications or grants with at least another UF investigator in that year (i.e., isolates are removed); and E is the set of collaborative ties (as defined above) among researchers in V. As is normally the case with publication and grant data, this is a weighted network, with the tie weight $e_{t,i,j}$ indicating the number of publications and grants on which researchers i and j collaborated in year t; and a longitudinal network including three waves (2013, 2014, 2015). Table 1 reports descriptive statistics on $G_t(V, E)$ over time. The number of nodes and edges in the network decreases slightly over the years. While density remains approximately constant, the higher number of components and lower proportion of nodes in the giant component reveal a slightly decreasing cohesion between 2013 and 2015 in publication and grant collaborations.

We briefly summarize here the procedure used to identify research communities in $G_t(V, E)$.[2] Intuitively, we think of a research community as a group of researchers that collaborate more closely and consistently with each other over the years, and more sparsely with other researchers. In the first step of the procedure, for each year, researchers are partitioned into collaborative subgroups using the Louvain community detection algorithm (Blondel et al. 2008). The algorithm detects communities or clusters of actors who densely interact with each other, but maintain sparser connections with the rest of the network. The Louvain procedure was specifically designed for large-scale networks with thousands of nodes (Fortunato, 2010), similar to the networks in our data. Furthermore, the procedure has been shown to detect network communities exposed to common spillover effects (e.g., in the diffusion of ideas or technologies), similar to our application (Eckles et al. 2017). Table 1 reports the main characteristics of the cluster partitions identified by the Louvain algorithm in $G_t(V, E)$ over 2013-2015. About 300 collaborative subgroups (clusters) are consistently identified over the years, with a stable distribution of subgroup size (subgroup average size

[2] More details are provided by Leone Sciabolazza et al. (2017).

ranging between 11.26 and 15.76 nodes). The algorithm returns a meaningful partition of nodes into internally cohesive and externally separate subgroups, as indicated by the high values of the modularity index (around 0.85).[3]

The extraction of collaborative subgroups in the first step produces three partitions of researchers into subgroups, one for each year in 2013-2015. Over this three-year period, two researchers can be co-members of the same subgroup, or belong to different subgroups. In the second step of the procedure, we represent patterns of subgroup co-membership among researchers constructing four co-membership networks, indicated as $G_{nc,1}$, $G_{nc,2}$, $G_{in,1}$, $G_{in,2}$. In networks $G_{nc,1}$ and $G_{nc,2}$, two researchers are connected if they were members of the same collaborative subgroup in two (*non-consecutive*) or three different years of the network $G_t(V, E)$, respectively (e.g., in G_{2013} and G_{2015}). In $G_{in,1}$ and $G_{in,2}$ a stricter definition of co-membership is applied, and two researchers are connected if they were members of the same *inter-temporal* collaborative subgroup respectively in two or three *consecutive* years (e.g., in G_{2013} and G_{2014}). An *inter-temporal* collaborative subgroup is defined as a temporal sequence of three collaborative subgroups (the first existing in 2013, the second in 2014, and the third in 2015) that include approximately the same researchers consistently over the years.[4] The four different co-membership networks correspond to weaker or stronger definitions of co-membership: two researchers may be connected in a "weaker" co-membership network (e.g., in $G_{nc,2}$) but not in a "stronger" one (e.g., in $G_{in,2}$).

The final research communities are extracted by applying the Louvain algorithm again to each of the four co-membership networks. The result is four sets of research communities, one for each co-membership network. The algorithm retrieves a meaningful community structure in all the co-membership networks, with substantially high values of modularity (0.40 - 0.65, Table 2). Between 49 and 189 research communities are identified in the university, with average size (number of scientists in the community) ranging between 6.44 and 23.16. Communities in the same set are mutually exclusive (they cannot include the same researcher), but communities in different sets can overlap. A research community identified by this method can be interpreted as a group of researchers who tend to work mostly with each other and to have common collaborators over the years, while interacting more rarely with other communities. The four sets include research communities resulting from increasingly strict definitions of co-membership and "stable collaboration", with the $G_{nc,1}$ set corresponding to the weakest definition, $G_{in,2}$ corresponding to the strongest, and $G_{nc,2}$ and $G_{in,1}$being in the middle.

Because the intervention was to be conducted as a pilot funding program for translational research in health, the intervention target population had to be delimited to research communities that included at least one researcher affiliated with an institute, department, or college in the health sciences. We refer to these as "health communities" or "health researchers" in the following. The main features of health communities are presented in Table 2. Health researchers represent more than 35% of the entire population of UF scientists involved in a detected research community. They are embedded in at least 43% of the research communities in the University, with these communities overall including more than 48% of all researchers.

The four co-membership networks produce research communities in different numbers and with different average sizes, resulting in varying productivity levels. [5] However, health communities in all sets are characterized by a similar demographic composition (on average community members are 50 years old, and at least 30% of them are female, Table 2), and by a consistently high density of interactions (46% to 67% on average), calculated as the proportion of dyads in one community that are connected in at least one year in $G_t(V, E)$. Moreover, health research communities show a consistently high rate of interdisciplinary collaboration: the average probability that two investigators in the same community are affiliated with different colleges (i.e., the average generalized variance of college affiliation in a community) is between 36% and 44%. This indicates that a high number of health researchers are likely to work with scientists from different disciplinary backgrounds and academic circles, potentially generating innovative collaborations. Most interdisciplinary collaborations involve co-authoring a paper, approximately one third are established for co-participating in a grant, and between 5 and 10% are created for both reasons. As expected, most members of health research communities are faculty (c.a. 80%), while a minority (c.a. 20%) is composed by staff (e.g., technicians, postdocs, etc.).

Table 3 examines the relationship among demographic composition, size, and rate of interdisciplinary collaborations in health research communities. Community size (A) and college diversity (B) are regressed on average investigator age and the percentage of women in a community (also including community size as a control in regression B). While community size does not significantly change at different levels of average age, it shows a negative and significant association with the percentage of women in a community in three out of four sets ($G_{nc,1}$, $G_{nc,2}$, $G_{in,1}$). By contrast, the degree of interdisciplinarity is not significant correlated with age or gender composition of communities in most cases, but it is always positively and significantly related to community size. These findings suggest three common patterns in all research communities. First, researchers of different ages are evenly distributed in communities of all sizes. Second, the percentage of women tends to be higher in small communities than in large ones. Third, large communities tend to be more diverse in disciplinary backgrounds (as proxied by colleges).

Table 4 shows the top 5 most represented colleges in each set of communities, and the percentage of researchers affiliated with them. Most researchers in each set of communities are affiliated with colleges in the health sciences (Medicine, Public Health and Health Professions, Dentistry), particularly in Medicine. However, three other colleges are almost always represented in this list: Engineering, Liberal Arts and Sciences, and Agricultural and Life Sciences. This finding is central to our study for two reasons. First, consistent with the high college diversity in the communities (see Table 2), it confirms the inclination of health communities toward interdisciplinary collaborations bridging different academic colleges, a central feature of interest in our population. Second, it provides additional evidence that broad institutional or administrative classifications such as colleges are unlikely to fully capture the patterns of actual scientific collaboration at a university, since most collaborative communities cut across different administrative entities (colleges, departments, schools, etc.). By contrast, network methods can identify research communities as they emerge from actual collaborative behavior (Leone Sciabolazza et al. 2017). This is a major reason why we consider research communities (rather than, for example, departments or colleges) to be the most appropriate targets and units of analysis for our intervention program.

[3] Modularity theoretically ranges between -1 and 1, with lower values reflecting poor community structure (many between-subgroup edges and few within-subgroup edges), and values closer to 1 indicating good community structure (few between-subgroup edges and many within-subgroup edges). In practice, modularity values above 0.3 are commonly considered as good indicators of meaningful community structure, with significant within-subgroup cohesion and between-subgroup separation (Clauset et al. 2004).

[4] Inter-temporal subgroups are identified by maximizing the Jaccard Index over all pairs of collaborative subgroups from two consecutive years (Leone Sciabolazza et al. 2017).

[5] Consistent with existing literature, scientific productivity is measured as the total number of published papers (Bosquet & Combes, 2013; Cainelli et al., 2015; Fafchamps et al., 2010; Hollis, 2001) and awarded grants (De Stefano et al., 2013) in 2013-2015. The number of per-capita published papers and awarded grants in one year (i.e., $\frac{average\ productivity}{3*(average\ community\ size)}$) is on average always

Table 2
Characteristics of co-membership networks and health research communities.

	$G_{nc,1}$	$G_{nc,2}$	$G_{in,1}$	$G_{in,2}$
Total number of research communities	189	126	109	49
Average size (SD) of research communities	23.16 (52.70)	6.44 (10.25)	15.33 (29.14)	10.91 (18.62)
Modularity of Louvain partition	0.65	0.59	0.63	0.40
Number of researchers affiliated with health institutes, departments, colleges (Percentage over total number of researchers in any research community)	429 (35%)	1084 (37%)	591 (35%)	206 (38%)
Number of health research communities (Percentage over total number of research communities)	94 (49%)	58 (46%)	47 (43%)	28 (57%)
Number of researchers in health research communities (Percentage over total number of researchers in any research community)	585 (48%)	2215 (75%)	1033 (62%)	287 (53%)
Size of health research communities: Mean (standard deviation)	6.22 (4.81)	38.18 (61.40)	21.90 (24.16)	10.20 (6.10)
Productivity of health research communities: Mean (standard deviation)	205 (228)	603 (1056)	471 (579)	372 (308)
Density of health research communities in $G_t(V, E)$: Mean (standard deviation)	0.67 (0.24)	0.40 (0.35)	0.46 (0.30)	0.59 (0.26)
College diversity in health research communities: Mean (standard deviation)	0.36 (0.26)	0.40 (0.28)	0.44 (0.27)	0.42 (0.24)
Density of interdisciplinary collaborations in health research communities: Mean (standard deviation)	0.46 (0.39)	0.37 (0.29)	0.39 (0.29)	0.48 (0.34)
Percentage of interdisciplinary collaborations: Mean (standard deviation)				
- Co-authoring a paper	45.03% (44.78%)	67.91% (32.12%)	57.79% (36.51%)	58.52% (39.94%)
- Co-participating in a grant	33.19% (42.26%)	23.02% (30.53%)	29.55% (34.03%)	19.03% (29.29%)
- Both	10.88% (15.68%)	4.53% (7.31%)	6.32% (9.89%)	11.22% (12.49%)
Researcher age in health research communities: Mean (standard deviation)	51.05 (12.08)	47.87 (12.49)	48.41 (12.37)	50.38 (11.87)
Percentage of female researchers in health research communities: Mean (standard deviation)	32% (47%)	34% (49%)	34% (48%)	30% (46%)
Percentage of researchers who are faculty in health research communities: Mean (standard deviation)	85.30% (27.57%)	77.60% (24.39%)	84.31% (19.98%)	88.25% (17.12%)
Percentage of researchers who are not faculty in health research communities: Mean (standard deviation)	14.69% (27.57%)	22.39% (24.39%)	15.56% (19.98%)	11.75 (17.12%)

Community *density*: proportion of connected pairs over all pairs in the community network. Community *productivity*: total number of published papers and awarded grants by researchers in the community (2013-2015). College *diversity* in community: generalized variance of colleges represented in the community. *Interdisciplinary* collaborations are collaborations between researchers affiliated with different colleges. *Density* of interdisciplinary collaborations: proportion of interdisciplinary collaborations over all connected and unconnected pairs in the community network. *Percentage* of interdisciplinary collaborations: number of interdisciplinary collaborations over all collaborations in the community network.

Table 3
The relation of investigators' gender and age with size and college diversity in Health research communities.

	(A) Dep. Variable: community size				(B) Dep. Variable: college diversity in community			
Partition	$G_{nc,1}$	$G_{nc,2}$	$G_{in,1}$	$G_{in,2}$	$G_{nc,1}$	$G_{nc,2}$	$G_{in,1}$	$G_{in,2}$
Age	−0.0112 (0.0784)	0.4873 (1.6117)	−0.0150 (0.6558)	−0.0439 (0.2596)	−0.0014 (0.0043)	0.0072 (0.0068)	0.0110 (0.0068)	−0.0274** (0.0081)
% Female	−0.0500** (0.0159)	−0.6443** (0.3189)	−0.2468* (0.1329)	−0.0527 (0.0550)	0.0740 (0.0919)	−0.2161 (0.1397)	0.1402 (0.1432)	−0.2420 (0.1758)
Community size	-	-	-	-	0.0186** (0.0057)	0.0019** (0.0006)	0.0053** (0.0016)	0.0130** (0.0063)
Num. Obs.	94	58	47	28	94	58	47	28
R2	0.0981	0.0770	0.0758	0.0357	0.1055	0.2700	0.2496	0.4137

Note: OLS coefficients (standard errors). *, **, *** indicate statistical significance at the 10, 5 and 1 percent level. An intercept is included. *Community size*: number of investigators in research community. *Community college diversity*: generalized variance of colleges in research community. *Age*: average age of investigators in research community. *% Female:* percentage female investigators in research community (multiplied by 100 in regression A).

Step 2: selection of potential collaborators within treatment communities

The overarching goal of the intervention is to foster collaboration and teamwork in research communities by creating a new collaborative relation. Therefore, in a given research community, the intervention should ideally create the link that maximally increases collaboration. This could be thought of as a problem of optimization in which we seek the missing link that would maximally increase a measure of collaboration and teamwork in a research community. We measure the overall level of collaboration in a scientific group using network cohesion indices. The cohesion of a network is a central and long-studied concept in social network analysis, which encompasses different dimensions (Borgatti et al. 2013; White & Harary, 2001). We use a standard, inverse measure of network cohesion, namely, the average geodesic distance between nodes (Borgatti & Everett, 2006; Freeman, 1979):

$$F = \frac{\sum_{i \neq j} d_{ij}}{N(N-1)} \tag{1}$$

where i and j are two nodes (researchers) in the subnetwork of a treatment community; d_{ij} is the geodesic distance between them in the subnetwork; and N is the total number of nodes in the community. F is an inverse measure of cohesion (or a measure of fragmentation), reflecting the idea that a network is more cohesive if its nodes are overall closer to each other in terms of geodesic distance. Lower levels of F characterize a network in which, by virtue of shorter geodesic paths among nodes, the diffusion of information, or any other resource flowing through the links, occurs more easily, quickly, and thoroughly. The underlying assumption is that network links are channels through which information and other resources pass from a node to another, such that longer geodesic distances are detrimental to network functioning. Research communities that are more cohesive according to F

Table 4
Top 5 most represented colleges in each set of health communities.

Co-membership network	College	College in health sciences	% affiliated researchers
$G_{nc,1}$	Medicine	Yes	55%
	Public Health and Health professions	Yes	7%
	Liberal Arts and Sciences	No	5%
	Dentistry	Yes	5%
	Engineering	No	4%
$G_{nc,2}$	Medicine	Yes	37%
	Engineering	No	10%
	Liberal Arts and Sciences	No	8%
	Agricultural and Life Sciences	No	7%
	Public Health and Health professions	Yes	4%
$G_{in,1}$	Medicine	Yes	42%
	Engineering	No	11%
	Public Health and Health professions	Yes	6%
	Liberal Arts and Sciences	No	5%
	Agricultural and Life Sciences	No	4%
$G_{in,2}$	Medicine	Yes	50%
	Public Health and Health professions	Yes	8%
	Engineering	No	5%
	Liberal Arts and Sciences	No	5%
	Dentistry	Yes	5%

% affiliated researchers is calculated over all researchers in a health community.

are also more similar to small-world networks, with a lower overall path length among nodes (Watts & Strogatz, 1998). This notion of cohesion is particularly appropriate to our application, and consistent with our view of collaboration and teamwork: we consider a scientific community as characterized by an overall higher level of collaboration (i.e., more cohesion according to F), when information, knowledge and scientific ideas can spread more easily and quickly among scientists throughout the community.

The measure F in Eq. 1 is used as an objective function in a fashion similar to Borgatti (2005) and Valente & Fujimoto (2010), where link deletion algorithms are used to assess the impact of one node, incident to the links, in decreasing the level of cohesion in a network. However, we are interested in adding missing links, not removing existing ones: we use F not to assess the importance of an *existing* link by measuring the cohesion *decrease* produced by its deletion; but to assess the importance of a *missing* link by measuring the cohesion *increase* produced by its addition. Thus, we first calculate F on the existing network of a scientific community. Then, for each pair of unconnected researchers in the community network, we simulate a new network in which that pair is connected, and recalculate F on the new network. Finally, we identify the pair of unconnected researchers whose newly created collaboration would *maximally decrease* F (i.e., maximally increase cohesion). More formally, let W be the set of all possible node pairs in $G_t(V, E)$, and let $U = \mathrm{W} - \mathrm{E}$ be the set of node pairs that are unconnected in the network. For each pair of nodes (i, j) that belongs to U and to the treatment community, we calculate $F_{ij}^{'}$, the average geodesic distance of the (simulated) community network in which i and j are actually connected. We then calculate the decrease in network fragmentation produced by the (i, j) link addition:

$$\Delta F_{ij} = F_{ij}^{'} - F \tag{2}$$

ΔF_{ij} is negative or null (the addition of a link cannot increase average geodesic distance: $F_{ij}^{'} \leq F$), with values close to zero indicating no fragmentation decrease. We select the pair (i^*, j^*) that minimizes ΔF_{ij}, and use the intervention to create an actual collaborative link between researchers i^* and j^* in the community.

In addition to being theoretically motivated as a problem of maximization of network cohesion, this procedure had another interesting advantage in our application. As shown in section 5, the minimization of ΔF_{ij} tended to identify pairs of potential collaborators (i^*, j^*) separated by higher geodesic distances, characterized by lower degree centrality (number of existing collaborations), and located in the fringes of the community network. Thus, our criterion tended to select missing links that were less likely to form naturally, because they connected scientists who were farther apart in the community network, therefore less likely to naturally meet and interact with each other in the workplace. Furthermore, the criterion tended to select individuals who were in peripheral areas of the community network, rather than in prominent roles and central positions in their respective scientific communities.[6] Very central scientists, with a high number of existing collaborations and a high level of involvement and power in a scientific group, were less likely to be selected by this criterion. In other words, this procedure was more likely to select individuals who, similar to the bridging actors described by Valente & Fujimoto (2010), were more inclined to participate in the intervention, because they had fewer existing collaborations, more time to invest in new teamwork, and less interest in preserving the status quo and current configuration of the research community. These characteristics are also consistent with Krackhardt (2001) suggestions on how to leverage network structural properties to foster exchanges of ideas and controversial innovations.

Step 3: creation of the missing link

The missing link was created by means of a UF CTSI limited submission pilot funding program, and by offering professional development funding to the two researchers in each selected pair if they submitted a letter of intent to the pilot program. First, we used the measure ΔF_{ij} (Eq. 2) to rank unconnected pairs of researchers in each treatment community, from the pairs with lowest ΔF_{ij} (those whose new collaboration would maximally increase community cohesion) to the pairs with highest ΔF_{ij}. The pair with lowest ΔF_{ij} was contacted via email and asked to participate. The email included a link to a visualization of the community research network, highlighting the nodes corresponding to the two invited researchers.

The email offered two incentives for the two investigators to participate in the intervention: (1) professional development funding ($1500 to each of the two potential collaborators in the pair) in exchange for taking part in an introductory meeting and submitting a letter of intent for the pilot program; and (2) expedited review and

[6] However, scientists who are peripheral in communities in the university network, might actually be more central in networks connecting different communities or institutions. In future projects we aim to replicate this intervention approach by considering communities that span different universities.

Table 5
Matching criteria for treatment and control groups.

Criterion type	Measure	Matching criterion
Average productivity	Average number of peer reviewed papers and awarded grants in 2013-2015	Max 2 SD between pairs
Network composition	College modal value	Equal
	College generalized variance	Equal
Network structure	Average degree centrality	Max 2 SD between pairs
	Size	Max 33% difference between pairs
No interference	Geodesic distance	The minimum geodesic distance between treatment and control communities in one pair is 2 or greater

special funding set aside for their pilot proposal (up to $25,000 per proposal) after successful peer review of the letter of intent and the full proposal. If both researchers agreed to participate, they were asked to take an eligibility survey to ascertain that they had indeed not collaborated on a publication or grant in the past three years. If either of the two researchers refused to participate or turned out not to be eligible, the pair with the next lowest value of ΔF_{ij} in the community was invited to participate.

Once the pair was successfully recruited, the intervention comprised the following phases:

i The two potential collaborators were asked to participate in a face-to-face meeting in which they were introduced to each other, and the design and purpose of the intervention was presented to them.
ii The pair was asked to submit a joint letter of intent within one month from the introductory meeting, to outline a proposal for the UF CTSI limited submission pilot program.
iii Upon submission of the letter of intent, each researcher in the pair received a compensation of $1500 in professional development funding that could be used for conference participation, research travel, or other professional development activities.
iv The submitted letters of intent we examined through standard peer review. If their letter of intent was selected, the pair was asked to submit a full proposal to the UF CTSI limited submission pilot program within three months from selection notification.
v The full proposals were peer-reviewed, and if selected were funded for up to $25,000.

Definition of the experimental setup and pairing of treatment and control communities

The impact of collaborations on team-level scientific productivity can be tested by evaluating the performance of a research community after a new collaboration is exogenously activated in it. The productivity change in the target communities can be assessed by comparing the performance of a treatment group, which includes the communities receiving the treatment (here, the creation of a new collaboration), with the performance of a control group, which is composed by a set of communities purposely designed to perform a counterfactual analysis. The statistical framework for this comparison is provided by the Rubin Causal Model (RCM) (Rubin 1973, 1974), a well-established approach to the analysis of causal effects in observational studies (Imbens & Wooldridge, 2009).

The essential component of the RCM is the definition of two potential outcomes. Ideally, the causal effect of a policy or intervention could be determined by measuring the outcome of the same unit in two “alternative futures”, in which the unit (here, the research community) receives different levels of exposure to the policy treatment. However, once the treatment is applied to a unit, only one of the alternative futures can be observed. To address this problem, the evaluation is conducted by comparing so-called potential outcomes: pairs of outcomes defined for the same unit given different levels of exposure to the treatment. As is standard in causal inference methods, given any treatment assigned to a group, the missing outcome associated with an alternate treatment is implicitly imputed by using a control group. The two potential outcomes are then compared to assess the impact of the intervention (for additional details the reader may refer to Imbens & Rubin, 2015). In this context, we define two potential outcomes: (1) The average individual productivity of the members of the research community when a new missing link is exogenously created by our intervention; (2) The average individual productivity in the research community when no intervention has taken place.

Different approaches exist to create treatment and control groups (Athey & Imbens, 2016). To limit costs, our intervention was designed as a pairwise randomized experiment, where each unit (here, a health research community) is paired with another unit with similar characteristics. One of the two units in the pair is then randomly extracted to be assigned to the treatment group, with the other being assigned to the control group. This mechanism improves the efficiency of the design by disallowing assignments that are likely to be uninformative about the treatment effect of interest, a problem that is usually addressed by increasing sample size, which was not a viable a solution in our case (Athey & Imbens 2016).

Table 5 summarizes the matching criteria we adopted to pair treatment and control communities. Following suggestions by Arpino et al. (2017), we pair two communities as treatment and control units if these are similar in terms of the individual characteristics of their members (network composition), and the structural characteristics of their collaborative ties (network structure). We consider two communities to have a comparable network composition if they have a similar distribution (modal value and generalized variance) of disciplinary affiliations, proxied by college. Communities in which most researchers belong to the same college (modal value of college affiliation) are more likely to show similar work practices and collaborative styles. Communities with similar degrees of college diversity (generalized variance of college affiliation) show a similar level of interdisciplinary collaborations. This leads to the following matching rules: the paired communities must have an equal value of (1) modal college affiliation and (2) college generalized variance.

As for network structure, the goal is to ensure that investigators in two paired communities are exposed to a similar collaborative environment. Hence, we regard two communities as comparable if they have a similar size (number of researchers), and if their members have on average the same number of collaborators within the community (average degree centrality). However, we do not seek exact matching for these measures to avoid too restrictive paring criteria, which would excessively limit the number of eligible pairs. Rather, we adopt a variation of the pairwise mechanism known as caliper matching (Cochran and Rubin, 1973), in which the observational units i and j are considered to match if $|P_i - P_j| < \varepsilon$, where ε is a pre-specified tolerance criterion (caliper) and P_i is the relevant measure for unit i. Thus, we adopt the following additional matching rules: (1) the paired communities must exhibit less than 33% difference in size,[7] and (2) the

[7] If P_i is the size of community i, this matching rule requires that $|P_i - P_j| < \frac{P_i}{3}$.

difference in average degree centrality between the paired communities must be less than two standard deviations of the degree distribution in community *i*.

Finally, to ensure that the treatment and control communities have comparable productivity, we consider the number of publications and awarded grants in the two communities over 2013-2015,[8] and adopt an additional matching rule: the difference between the average productivity in paired communities *i* and *j* must be less than or equal to two standard deviations in *i*'s distribution of individual productivity.

The last issue to take into account in RCM designs is the possibility of interference, that is, of interaction between treatment and control units. The impact of the intervention can only be correctly estimated if there is no interference: control communities do not benefit from the effects of the intervention on treatment communities. When individuals can be partitioned into groups, as in our case, it is often plausible to assume that interference occurs within groups but not across groups (Sobel 2006). Recent literature also leveraged the knowledge of network structure to minimize the risk of interference (Arpino et al., 2017; Coppock & Sircar, 2013; Eckles et al., 2017; Ugander et al. 2013). In this line of research, treatment units are selected in such a way that (1) they are separated by a pre-determined social distance from control units, measured as geodesic distance (length of the shortest network path), or (2) they belong to a network cluster where there are no control units (with network clusters being identified by a community-detection algorithm). We adopt the first approach and add one restriction criterion to our matching process: two communities can only be paired if no collaborations have taken place between their members in 2013-2015. In other words, the minimum geodesic distance between researchers in a treatment community and in its respective control community must be 2 or higher in $G_t(V, E)$ ($t \in${2013, 2014, 2015}). The idea underlying this criterion is that past collaborations reveal investigators' collaboration preferences, and if two researchers have not interacted with each other and have consistently shared no collaborators for 3 years, we assume that there are negligible chances that they will start to work together during our intervention, potentially biasing its evaluation. The Appendix A provides a detailed description of the algorithm used to identify treatment and control communities based on the matching criteria in Table 5.

Potential sources of bias in the experimental design

In this subsection, we examine the robustness of the presented counterfactual design, verifying that the match between treatment and control communities is satisfactory for comparisons and warrants a causal interpretation of future results. The ideal way to ensure correct matches is by introducing additional (pre-treatment) variables in the model, and reducing caliper criteria, even to the extent to saturate the model: i.e., adopting matching criteria that include all observed variables and possible product-terms, with all calipers being set to zero. This would ensure perfect matching between treatment and control units, ruling out any potential bias in the counterfactual design. However, while this approach is possible in principle, it is usually not feasible in practice. This is because the design of a matching strategy is necessarily driven by the nature of the population (Imbens & Rubin 2015). In this study, the intervention targets a relatively small population of research communities, where pre-treatment variables (e.g., those listed in Table 5) take many distinct values. As a result, even a nearly-saturated model, including a high number of pre-treatment variables, would generate a substantial number of unpaired communities in this design. Matching can be even more problematic for particularly small communities, including less than ten nodes, for which appropriate counterfactuals are unlikely to be found with criteria including many pre-treatment variables.

Another reason why we limit the number of pre-treatment variables for matching criteria is to preserve the feasibility and replicability of our intervention and evaluation design. This study aims to propose an approach that is reasonably feasible given commonly available data on academic workforce at research universities. Data on publications and grants are generally easier to obtain from existing, and relatively more open, university databases. By contrast, individual-level information on researchers' sociodemographic characteristics, job positions and professional achievements (e.g., age and sex, job titles, academic awards), is typically unavailable, more protected, or too expensive to obtain in university settings, requiring manual mining of thousands of CVs or costly surveys with thousands of respondents. For these reasons, we adopted a matching strategy based on a limited set of covariates that are easy to collect and likely to drive productivity in research communities; while at the same time using the caliper approach to obtain a sufficient number of paired communities to involve in the program.

It is also important to note that variations in caliper values have a relatively small impact on the number of matched communities in our application. Fig. 1, for example, shows the effect of network structure calipers. In panel (a), all else being equal, setting the community size caliper to zero decreases the number of paired communities by only 4 observations (from 15 to 11 pairs). The average degree caliper has a stronger effect, reducing the number of matched pairs to 4 when set to 0 (panel b). The reason is that average degree centrality is not an integer value, and finding a perfect match is unlikely (for example, even a difference in the third decimal digit would prevent a match). Unsurprisingly, even a very small increase in the caliper value (from 0 to 0.5 standard deviations of the degree centrality distribution) is enough to increase the number of matches by almost 300% (from 4 to 11 matches). Further expanding the caliper from 0.5 to 2 standard deviations increases the number of eligible community pairs by only 4 observations (from 11 to 15 pairs). In our design, caliper values were selected based on a data-driven approach that considered the trade-off between two goals: pairing communities that are as similar and comparable as possible; and obtaining a number of matches sufficiently high to make the intervention possible. On the other hand, while appropriate in our application, the calipers adopted here are not proposed as standard, universally acceptable values to be replicated in future research. Researchers interested in replicating this study may consider different matching strategies and caliper values, depending on the nature of their population, the practical constraints of their intervention, and the impact of caliper variation on matches in their particular data.

Even if calipers might not represent a major issue, one might still doubt that our matching strategy leads to an effective counterfactual design, with comparisons between treatment and control units potentially biasing evaluation results. An approach to addressing this concern is to match units on a set of significant (pre-treatment) variables (such as the ones in Table 5), so that other relevant characteristics of the units become ignorable (Imbens & Rubin 2015). This approach is reliable if the counterfactual design results in similarity between treatment and control units even on meaningful characteristics that were not explicitly included in the matching criteria. Section 5.1 shows that this is the case for our matching strategy.

A comparison between paired communities by pre-treatment variables (Tables 6–7) allows us to assume that the assignment mechanism to treatment and control groups is exchangeable, that is, a community is not assigned to treatment or control because of its characteristics. Moreover, by imposing a pre-determined social distance between

[8] The idea here is that grants represent an important gateway for establishing research collaborations and they are a significant proxy for future publications. This is consistent with existing evidence showing that investigators who collaborate on a research grant will co-author at least one publication within 5 years (Jacob & Lefgren, 2011). Therefore, by combining number of publications and number of awarded grants, we aim to obtain a measure of productivity which includes academic outcomes in the long term, and to adopt a more restrictive criterion when matching treatment and control communities.

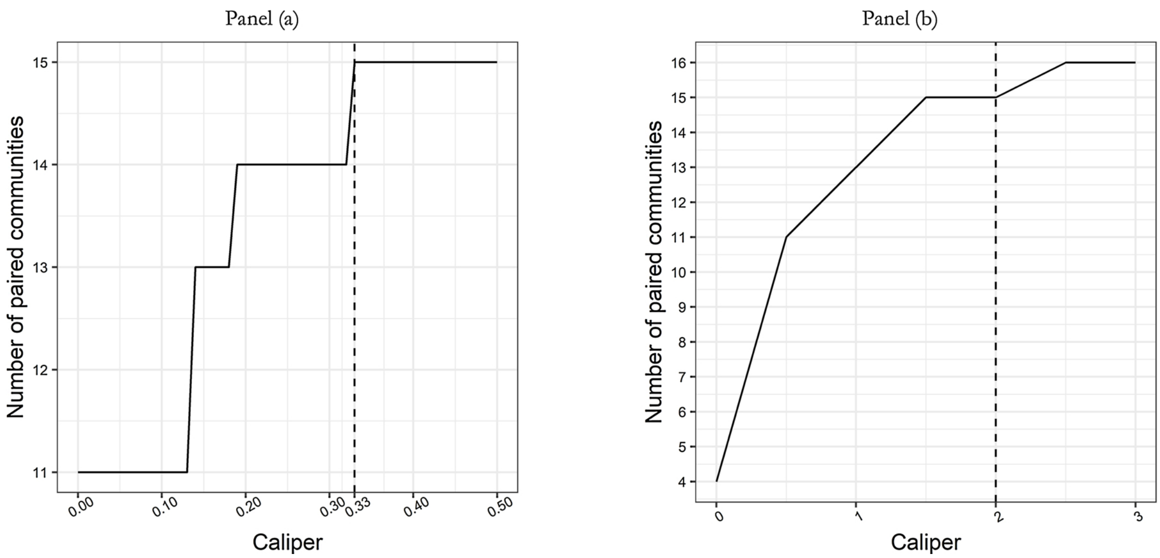

Fig. 1. The impact of network structure calipers on the number of paired communities.
Note: The x axis in panel (a) reports the caliper value used to match communities' size (max percentage difference between pairs). The x axis in panel (b) displays the caliper value used to match communities' average degree centrality (max difference between pairs in standard deviations). Y axis indicates the number of paired communities after step 2 is completed for any given value of caliper used (see Section 4.2). The dotted line shows the caliper value set in the experiment design (see Table 5)

Table 6
Main characteristics of treatment and control research communities.

Co-membership network	Community id	Avg productivity Treatment/Control	Network composition		Network structure	
			College modal value Treatment/Control	College generalized variance Treatment/Control	Average degree centrality Treatment/Control	Size Treatment/Control
$G_{nc,1}$	1	33.11 / 16.25	Medicine / Medicine	0.35 / 0.61	7.11 / 4.83	9 / 12
	2	10.25 / 12.43	Medicine / Medicine	0.00 / 0.00	2.50 / 2.00	4 / 7
$G_{nc,2}$	3	9.90 / 4.29	Medicine / Medicine	0.32 / 0.28	4.40 / 2.86	10 / 7
$G_{in,1}$	4	11.19 / 15.73	Medicine / Medicine	0.71 / 0.58	5.12 / 4.97	43 / 37
	5	6.50 / 4.67	Medicine / Medicine	0.38 / 0.44	2.50 / 1.33	4 / 3

Table 7
Test on excluded matching criteria.

		Community ID				
		(1)	(2)	(3)	(4)	(5)
Research Seniority	Age	−0.2081	−1.4412	0.5774	−1.5908	1.1921
		(0.8357)	(0.1838)	(0.5889)	(0.1285)	(0.2518)
	Years since Ph.D.	0.1548	−1.3845	−0.3974	−0.3739	0.8258
		(0.8774)	(0.2010)	(0.7199)	(0.7127)	(0.4232)
Seniority at university	Studied at UF (1 = Yes)	0.4348	0.0000	0.1094	3.5379*	0.0000
		(0.5096)	(1.0000)	(0.7409)	(0.0600)	(1.0000)
	Years working at UF	0.9436	−3.873***	0.3333	0.1270	0.5423
		(0.3483)	(0.0082)	(0.7544)	(0.9006)	(0.6005)
Job position	Academic Position (0 = Staff member, 1 = Faculty)	−2.7883*	0.0000	0.0243	0.0000	0.2277
		(0.0950)	(1.0000)	(0.8761)	(1.0000)	(0.6333)

Note: a t-test and p-value (in parentheses) is reported to assess statistical difference between treatment and control communities for the variables Age, Years since Ph.D. and Years working at UF. A chi-square test and p-value (in parentheses) is reported to assess statistical difference between treatment and control communities for the variables Studied at UF (1 = Yes) and *Academic Position* (0 = Staff member, 1 = Faculty). *, **, *** indicate statistical significance at the 1, 5, and 10 percent level.

paired communities, we rule out the risk that they interact with each other due to similar characteristics (i.e., homophily mechanisms), introducing interference in the experiment. This ensures that no confounding effects are at play in the experiment, and causal interpretation of intervention results is correct.

Results: Implementation of the intervention

The matching rules (Table 5) and algorithm (Appendix A) were used to identify intervention targets among the research communities detected in Step 1 of the program.[9] The method detected 15 pairs of

[9] The results of Step 1 are fully detailed by Leone Sciabolazza et al. (2017).

treatment and control communities in the health sciences, of which 2 pairs derived from network $G_{nc,1}$, 7 from $G_{nc,2}$, 5 from $G_{in,1}$, and 1 from $G_{in,2}$. While this was not explicitly included as a selection criterion, none of the 30 communities had any investigator in common. Consequently, we did not need to consider estimation problems that may arise from the presence of investigators embedded in different groups. In Step 2, in each of the 15 treatment communities we identified unconnected pairs of potential collaborators, whose connection would maximally increase community cohesion. In Step 3, 30 researchers (15 pairs, one for each treatment community) were contacted and offered to participate in the program. One third of them (5 pairs) agreed and were eligible to participate.[10] The five pairs took part in an introductory meeting and submitted a letter of intent for the limited submission pilot program. The letters of intent were peer-reviewed and three out of five were selected for the full proposal submission. Three full pilot proposals were received and, after successful peer-review, funded for $25,000 each.

The selection of collaborator pairs based on cohesion maximization within research communities resulted in meaningful matches, with the selected researchers being relatively close in the collaboration network in terms of geodesic distance. By constraining the selection of potential collaborators to researchers who belonged to the same community, we were able to link investigators with overlapping research interests, even if they had never directly worked together or were not affiliated with the same department, college or institute. As a result, each pair was able to identify a common topic of interest for collaborating on a research proposal:

1 The first pair (community 1), composed by faculty in the Department of Pharmacology and Therapeutics and in the Department of Pediatrics, submitted a pilot proposal on smoking cessation therapies during pregnancy and their association with maternal and infant health outcomes.
2 The second pair (community 2), composed by researchers in the Department of Medicine, submitted a proposal to develop new models for influenza infection for optimizing anti-influenza therapies.
3 The third pair (community 3), composed by researchers in the Department of Dentistry and in the Department of Engineering, submitted a proposal to improve reliability and lifetime prediction analyses of dental ceramic materials.
4 The fourth pair (community 4), composed by investigators from the Department of Pediatrics and the Department of Biomedical Engineering, submitted a proposal to study diabetes therapies based on immunotherapy and tolerogenic microparticle vaccine.
5 The fifth pair (community 5), composed by faculty from the Department of Pediatrics and the Department of Obstetrics & Gynecology, submitted a proposal to improve diabetic wound healing and reduce liver cancer burden.

Main features of selected communities

Table 6 shows the main characteristics of the treatment and control communities. Members of most treatment communities (4 out of 5) exhibit higher-than-average scientific productivity in 2013-2015. The average number of publications and grants per investigator in 2013-2015 is 5 on average for health communities, while it is approximately 10 for three of the selected communities, and 33 for one community, a value remarkably high when compared to the rest of the university. Only for one of the five target communities is scientific productivity in line with the overall population (6 publications/grants compared to the average of 5 in health communities). This misalignment risks to bias the internal validity of the experiment, since the sample is poorly representative of the entire population of research communities in the university. This is a common problem when using a pairwise matching mechanism, because its restrictive pairing criteria significantly reduce the number of eligible pairs in the population, generating a sample that may not be statistically representative. Future implementations of this intervention should involve a higher number of communities to select eligible units with less restrictive matching mechanisms (e.g., random assignment) and obtain a better representation of a population of research communities of interest. Obviously, a higher number of treatment communities is likely to make the intervention more costly, as incentives for participation (here, professional development funding) have to be offered to a higher number of potential collaborators.

As far as network structure is concerned, community characteristics are more heterogenous (Fig. 2). Two communities are only composed of 4 members. In this case, the intervention consists in connecting the pair of investigators who are not collaborating within the community. Two communities are composed of approximately 10 members, however the average number of collaborations (average degree centrality) is higher in one community than the other (7.11 compared to 4.40). This difference is consistent with differences in productivity, with the community with the highest average number of collaborations also being the most productive in our sample. The last community is the largest one, with 43 members and 5 collaborations per investigator on average.

With regard to network composition, consistent with our reference population: i) most researchers are affiliated with the College of Medicine in 4 out of 5 communities; ii) the probability of collaborations occurring between different colleges in a community (college generalized variance) is about one third or higher in 4 of the 5 treatment communities. Unsurprisingly, the largest community (43 members) is also the one with the highest college diversity (0.71) and potential for interdisciplinary collaborations. Only one of the five communities has all members in the same college.

Findings in Table 7 support the robustness of our counterfactual design by showing that differences are negligible between treatment and control communities on potentially relevant variables other than those used in the matching criteria. We conducted t-tests and chi-square tests comparing treatment and control communities on a set of variables which can possibly be related to productivity, but are not included in the matching criteria.[11] These variables included: i) *research seniority*, expressed in terms of age and number of years since Ph.D.; ii) *seniority at university*, measured by a dummy variable indicating if the researcher studied at UF or not, and by the number of years since the researcher was hired at UF; iii) *job position*, i.e., whether the researcher is faculty or staff member. The results show that differences between treatment and control communities are not statistically significant in almost all cases. Two communities (3 and 5) have no statistically significant differences from their controls on any variable. The other three communities show a significant difference with their respective controls only on one variable (one measure of *seniority at university* in two cases, and *job position* in one case).

Main features of selected pairs

Table 8 describes the main features of the pairs of potential collaborators recruited in the treatment communities. In three out of five cases, the intervention connects scientists from two different departments, and in two cases different colleges. Thus, in most cases the program seems effective at facilitating interdisciplinary collaborations

[10] While there were multiple reasons the remaining 10 pairs did not participate, a recurrent reason was that at least one person in the pair was already a very productive researcher and did not feel they needed more collaborators. This is a problem in that very successful and central researchers may become stuck in their central positions, and structurally remove opportunities for advancement of some nodes in their neighborhood.

[11] These variables were manually collected by mining CVs of researchers in treatment and control communities.

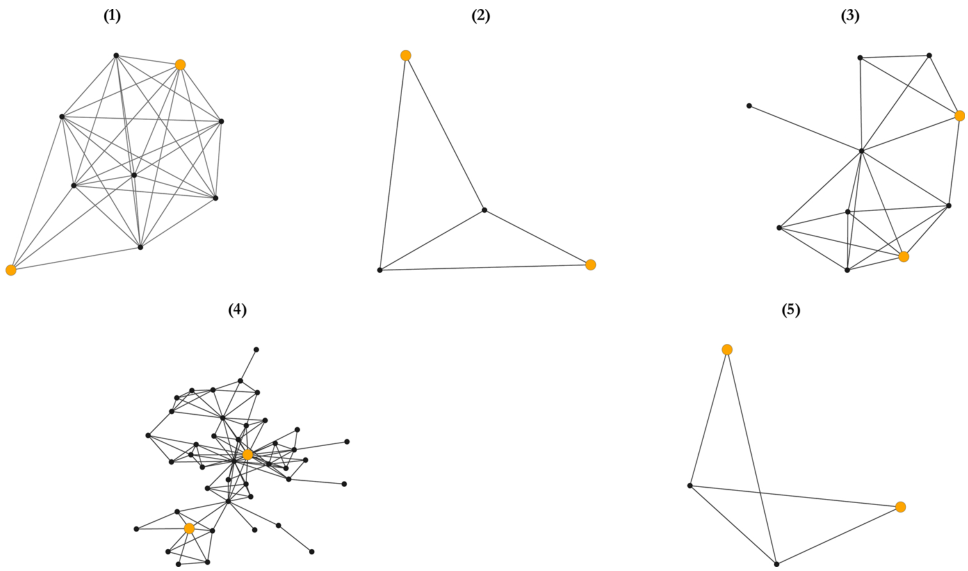

Fig. 2. Subnetworks of health research communities in the treatment group.
Note: each dot (node) represents an investigator. Two investigators are connected if they have collaborated (i.e. if they are linked in at least one year in G_t (V,E)). Orange nodes are the investigators selected for the intervention. Numbers are community ids. (For interpretation of the colour in this figure legend, the reader is referred to the web version of this article).

that cut across different administrative units. In terms of productivity, in four out of five pairs we find a significant difference in the number of the research outputs between the two potential collaborators. This is probably due to the fact that most of the recruited investigators are affiliated with different departments, which may imply different productivity standards.

Interestingly, most of the paired investigators have a similar number of other collaborators in the community (degree centrality), suggesting that most of them are equally involved in their research group. Six of ten selected researchers have a relatively low number of collaborators in the community (between 1 and 4), confirming that the intervention procedure tends to select investigators in peripheral areas of the community networks, presumably with less prominent roles and involvement in their respective community. In every pair, the two selected researchers have at least one collaborator in common within the community, i.e., a geodesic distance of two links in the community network. This characteristic, which is more likely when selecting pairs within the same research community, facilitates the overlap of research interests and mutual trust between potential collaborators. At the same time, paired investigators share no collaborators outside their community, a result suggesting that the community identification method used for this study was successful at detecting internally coherent and externally separate research groups.

Discussion: Challenges and expected outcomes

Challenges in experimental interventions in scientific networks

Three major challenges had to be addressed in the design of the intervention program presented in this paper. The first was the selection of potential collaborations that would be considered as acceptable by intervention participants: that is, the identification of pairs of investigators with compatible research interests. The second was the creation of appropriate treatment and control groups considering that observational units (research communities) are embedded in a social network. The third was the risk of interference between treatment and control group.

The first issue was addressed by conducting the intervention within groups of investigators (research communities) who are likely to share a set of common interests because they have consistently worked together or with common collaborators in the three years before the intervention. This approach led to meaningful pairings between potential collaborators, with target scientists being able to identify topics of common interest to start their collaborations. The second issue was tackled using the innovative framework provided by Arpino et al. (2017) for the construction of a robust counterfactual when observational units are embedded in a social network. This method consists in matching a treatment and control unit when they share individual characteristics, their peers present common features, and they are embedded in structurally similar networks of collaborations. The third issue was addressed by selecting pairs of treatment and control units located in different and distant areas of the scientific collaboration network. This led to the recruitment of scientists in the treatment group who were not likely to collaborate with colleagues in the control group during the time of the intervention, hence avoiding risks of interference across groups.

A different challenge is related to the accurate evaluation of a collaboration's causal effect on productivity when observation units are embedded in a social network. Different problems can arise when the impact of scientific collaborations is assessed using observational data, rather than experimental approaches (Fafchamps, 2015). A key problem is that two distinct causal mechanisms, which may explain a highly productive collaboration, cannot be distinguished in observational data (Manski, 1993). The first mechanism is based on homophily and social pressure: two researchers, who share similarly high productivity, begin a collaboration (homophily), and this creates pressures that drive them to continue working with one another and remain highly productive together. The second mechanism is based on influence and peer effects: collaborators influence each other, building up new skills, creating an efficient combination of expertise, and becoming more productive together than they are on their own (knowledge spillover). As shown by Advani & Malde (2014) and tested by Comola & Prina (2014), a crucial advantage of experimental approaches is that, by exogenously adding a link between two nodes that are otherwise unconnected, the homophily mechanism can be excluded: in this context the productivity of the pair before the collaboration is observed, and the output of their collaboration is only generated by their ability to work together. In this case, the output of the collaboration can be studied only as a function of i) investigators' common interests and related methods, and ii) the influence exerted by their community of collaborators. As a result, using network methods for experimental data (Arduini et al. 2019; Del Prete et al., 2019, Forastiere et al. 2017), one can discriminate between the social pressure and the peer effect mechanisms, and distinguish the impact of the intervention (i.e., of the creation of a new collaboration) from the impact of the investigators' peers.

Table 8
Main characteristics of selected pairs in treatment research communities.

Community id	Pair id	Department	College	Productivity	Degree centrality in the community	Geodesic distance between pairs	Number of shared collaborators within the community	Number of shared collaborators outside the community	Decrease in average geodesic distance
1	1	Pharmacology and Therapeutics	Medicine	20	7	2	4	0	2.50%
	2	Pediatrics	Medicine	12	4				
2	1	Medicine	Medicine	7	2	2	2	0	1.47%
	2	Medicine	Medicine	3	2				
3	1	Dentistry	Dentistry	5	1	2	2	0	12.50%
	2	Engineering	Engineering	7	2				
4	1	Pediatrics	Medicine	12	7	2	1	0	5.24%
	2	Biomedical Engineering	Engineering	49	18				
5	1	Pediatrics	Medicine	13	5	2	2	0	1.47%
	2	Obstetrics & Gynecology	Medicine	6	4				

Expected intervention outcomes

The intervention program presented here is expected to yield measurable results a year after its completion in two distinct domains, namely, scientific collaboration and productivity; and at three different levels, that is, for individual scientists, dyads, and whole research communities.

Scientific collaboration is operationalized in this study as co-authorship on publications and co-participation in externally funded research grants. Outcome measures in this domain will include the weighted degree of researchers in publication and grant networks (i.e., the number of collaborative projects a researcher is involved with), for the individual level; the number of publication and grant collaborations in a target pair, for the dyad level; and measures of collaborative cohesion, such as density of the collaboration network, for the community level. The new link created by the intervention program consists of a collaboration on a pilot, internally funded research grant. We hypothesize that this link will lead to further collaborations on publications and grants in the target dyads. We also hypothesize a cascading effect in which more collaboration in the target dyad encourages additional collaborations by each individual in the dyad, and by other individuals in the treatment community, for example due to transitivity and social circuit effects (Snijders 2011).

Scientific productivity is measured as the number of peer-reviewed articles and research grants, as well as the number of citations and amount of grant funding to adjust for impact. We hypothesize that the intervention will have both a direct and an indirect effect on productivity. First, the intervention program directly supported new research for the recruited individuals and dyads, which should result in higher productivity. Second, we expect the intervention to have an indirect, collaboration-induced effect on productivity: based on existing studies on the benefits of team science and interdisciplinarity (Lee & Bozeman 2005, Wuchty et al. 2007, Leahey 2016), we hypothesize more collaboration to facilitate communication and cross-pollination of scientific ideas and opportunities both in the target dyads and in the broader treatment communities. The final result of this process should be overall higher scientific productivity and impact for individuals, target dyads, and treatment research communities.

Constrained network alteration for research funding

Traditional mechanisms for research funding, including intramural pilot or seed programs, can be regarded as a form of network induction (Vacca et al. 2015): a type of intervention that does not alter the existing social network, but stimulates interactions largely through existing network links. In a traditional research funding program, the funder advertises an open request for applications, soliciting submissions on more or less narrow research topics from any researcher or team in a given population. Although there may be team composition requirements (for example, the team may be required to include members from more than one university or more than one department), traditional programs let the applicant teams emerge from an *existing* network of professional relationships and scientific collaborations. As a result, scientists tend to apply with collaborators in their professional comfort zone, with largely overlapping background and expertise, often colleagues with whom they have already worked in the recent past. In this way, the existing structure of the scientific network heavily conditions the future evolution of science (Fortunato et al., 2018). Therefore, traditional research funding programs may not be the most effective way to fund truly novel collaborations that explore innovative combinations of ideas from distant disciplinary backgrounds. In particular, traditional programs cannot be used to conduct forms of network alteration: interventions that purposely change the existing scientific network, for example creating specific, new collaborative links that maximize the likelihood of positive outcomes or optimize desirable network properties. Inductive interventions may have a tendency to solidify existing structural boundaries over time, benefiting large research groups at the expense of newly emergent research that works across those boundaries.

Alternative funding mechanisms are being explored to support interdisciplinary research, including person-directed funding that does not involve peer review (Azoulay et al. 2011; Fortunato et al., 2018). However, innovative research funding programs that use tax dollars and public resources in radically novel ways are likely to encounter resistance and pushback from existing stakeholders, particularly if they imply direct selection of awardees and limit competitive applications and peer review. Our intervention program seeks a compromise between the benefits of network alteration and person-directed funding on the one hand, and the need to allocate research funding in ways that are competitive and politically acceptable on the other. While our program does directly select potential awardees to conduct network alteration, it also preserves competition and peer-review in the proposal selection process.

Existing literature on team science emphasizes the tension between the societal need to fund truly novel, interdisciplinary combinations of ideas from distant areas of science, on the one hand; and the individual need to collaborate with professionally close and familiar colleagues to limit teamwork costs and risks, on the other hand (Leahey, 2016). Our proposal seeks to reconcile this tension. While we create new collaborative links identified by a cohesion-optimization algorithm that tends to pair people from distant areas of a network, we also constrain the algorithm to search for pairings within the boundaries of established research communities. In this method of *constrained* network alteration, the potential collaborations we propose are novel and typically interdisciplinary, yet they involve researchers who are not too distant in the network and do not perceive each other as too unfamiliar and risky a collaborator. Although our algorithm could be applied to larger areas of the scientific network (for example, the network of the entire College of Medicine) to identify missing links that would increase overall network cohesion even more, this application would likely suggest potential collaborations that would be perceived as too costly and risky by individuals, as demonstrated by previous experiments (Vacca et al., 2015). At the same time, a key aspect of our intervention is that the group boundaries within which the algorithm operates and the new collaborations are selected, are not defined based on a priori institutional affiliations, but based on the actual collaborative interactions emerging from the data: we search for optimal pairings not within institutional departments or other administrative entities, but within research communities emerging from the actual collaboration networks in the university. This increases the likelihood of proposing truly interdisciplinary collaborations.

Conclusions

This paper presented a network intervention to foster scientific collaboration and productivity at a research university, and an experimental approach to evaluate the intervention's impact. The intervention consists in identifying research communities in a university's publication and grant collaboration network, and creating a new collaborative relation between two unconnected researchers in treatment communities. The new collaboration is created in such a way that the overall cohesion of the target community is maximally increased, fostering the diffusion of information, knowledge and scientific ideas in each treated scientific group. This intervention program draws on the notions of segmentation and alteration established in previous studies of network interventions. In line with suggestions from recent literature in the Science of Science (Fortunato et al. 2018), we also presented an experimental setup for rigorous evaluation of the program's impact, which is designed according to the Rubin Causal Model framework.

This project is innovative in three ways. First, it explores the application of well-established notions of network intervention to a new field, namely, the study of science and research policy. Second, it implements a program of network alteration, which is a relatively more costly, difficult and rare form of network intervention (Valente, 2012). Third, it proposes novel ways of designing a very popular type of research funding programs in academia, namely, intramural pilot or seed funding programs. While innovative, this intervention program was also grounded in well-established knowledge and practices in university research funding. It was managed by an academic research institute (the UF CTSI), which is similar to the many institutes externally and internally funded at US universities to promote interdisciplinarity in the biomedical and social sciences (Jacobs & Frickel, 2009). Furthermore, it was implemented as a novel, experimental version of an existing pilot funding program that was well-known in its university, and akin to the many seed or pilot programs adopted at other US research universities.

We hope that the experimental intervention presented in this article will stimulate further research on the use of network data and methods to inform policies on research and science. Key steps in this endeavor will be the evaluation of scientific network interventions in different academic contexts (e.g., public and private research universities of different sizes and in different geographical contexts); and the scaling-up of similar interventions to research consortia and networks that connect different universities and potentially different countries.

Sources of funding

Research reported in this publication was supported by the University of Florida Clinical and Translational Science Institute, which is supported in part by the NIH National Center for Advancing Translational Sciences under award number UL1TR001427.

Appendix A. Algorithm to Identify Treatment and Control Communities

The following algorithm was implemented to identify pairs of matching communities to assign to treatment and control. A matching pair is defined as a pair of communities (j, k) that matches according to the rules in Table 5. A unique match is defined as a matching pair (j, k) such that j and k are not matched with any other community (i.e., no other matching pair exists which involves j or k). A multiple match is defined as a set of multiple matching pairs that involve the same community, e.g. the pairs (j, k) and (j, l).

We start by assigning a random integer numeric ID to each community, $i \in \{1, 2, 3, \ldots, n\}$, where n is the total number of communities. We then create three sets: F is the final set of matched community pairs; T is the temporary set of matching community pairs; U is the set of all unmatched communities. At the beginning F and T are empty, and all community IDs are in U. The purpose of the algorithm is to store as many community pairs as possible in F.

The algorithm works as follows. Set $i = 1$ and conduct the following iteration:

1 Select community i in set U.
2 Search for all matches in U for community i based on rules in Table 5.
   a If no matches are found, then go to Step 3.
   b If multiple matches are found, then:
      i Store all of i's matching pairs in T. (For example, if i is found to match with both j and k, then store (i, j) and (i, k) in P.)
      ii Go to Step 3.
   c If a unique match is found between i and community j, then:
      i Store (i, j) in F.
      ii Remove i and j from U.
      iii Remove from T any matching pair involving i or j (potentially resulting from previous iterations).
      iv Check if a new unique match exists in T (following removal of pairs involving i or j). If so, go to Step 2c and apply it to new unique matching pair.

v Go to Step 3.

3 If $i = n$ stop. Else set i to next available integer in U and go to Step 1.

When the algorithm stops, if T still contains multiple matches, we use a maximization function to ensure that (i) each community appearing in T is matched with one and one only community, and (ii) the highest number of communities appearing in T has at least one match. The resulting community pairs are then stored in F. In our application, however, T was found empty at the end of the algorithm, so this operation was not necessary.

The algorithm was separately executed on each of the four sets of communities (resulting from each co-membership network). The algorithm produced 15 pairs of matching research communities (30 communities in total). In each of the 15 pairs, one of the two communities was randomly assigned to treatment and the other was assigned to control.

References

Acedo, F.J., Barroso, C., Casanueva, C., Galán, J.L., 2006. Co-authorship in management and organizational studies: An empirical and network analysis. Journal of Management Studies 43 (5), 957–983.
Advani, A., Malde, B., 2014. Empirical methods for networks data: Social effects, network formation and measurement error. IFS Working Papers 14 (34).
Aral, S., 2016. Networked Experiments. In: Bramoullé, Y., Galeotti, A., Rogers, B. (Eds.), The Oxford Handbook of the Economics of Networks, pp. 376–411.
Arduini, T., Patacchini, E., Rainone, E., 2019. Treatment Effects with Heterogeneous Externalities. Journal of Business Economic Statistics Forthcoming.
Arpino, B., De Benedictis, L., Mattei, A., 2017. Implementing propensity score matching with network data: the effect of the General Agreement on Tariffs and Trade on bilateral trade. Journal of the Royal Statistical Society: Series C (Applied Statistics) 66 (3), 537–554.
Athey, S., Imbens, G., 2016. The state of applied econometrics-causality and policy evaluation. ArXiv Preprint ArXiv:1607.00699.
Azoulay, P., Graff Zivin, J.S., Manso, G., 2011. Incentives and creativity: evidence from the academic life sciences. The RAND Journal of Economics 42 (3), 527–554.
Bikard, M., Murray, F., Gans, J.S., 2015. Exploring Trade-offs in the Organization of Scientific Work: Collaboration and Scientific Reward. Management Science 61 (7), 1473–1495.
Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 1–12 (P10008).
Borgatti, S.P., 2005. Centrality and network flow. Social Networks 27 (1), 55–71.
Borgatti, S.P., Everett, M.G., 2006. A graph-theoretic perspective on centrality. Social Networks 28 (4), 466–484.
Borgatti, S.P., Everett, M.G., Johnson, J.C., 2013. Characterizing whole networks. In: Johnson, J.C., Everett, M.G., Borgatti, S.P. (Eds.), Analyzing social networks. Sage Publications, pp. 149–162.
Bosquet, C., Combes, P.-P., 2013. Do large departments make academics more productive? Agglomeration and peer effects in research. CEPR Discussion Paper (DP9401).
Buller, D.B., Morrill, C., Taren, D., Aickin, M., Sennott-Miller, L., et al., 1999. Randomized Trial Testing the Effect of Peer Education at Increasing Fruit and Vegetable Intake. JNCI: Journal of the National Cancer Institute 91 (17), 1491–1500.
Cainelli, G., Maggioni, M.A., Uberti, T.E., de Felice, A., 2015. The strength of strong ties: How co-authorship affect productivity of academic economists? Scientometrics 102 (1), 673–699.
Chung, K.H., Cox, R.A.K., Kim, K.A., 2009. On the relation between intellectual collaboration and intellectual output: Evidence from the finance academe. The Quarterly Review of Economics and Finance 49 (3), 893–916.
Cialdini, R.B., Trost, M.R., 1998. Social influence: Social norms, conformity and compliance. In: In: Lindzey, G. (Ed.), The handbook of social psychology Vols. 1-2. pp. 151–192.
Cochran, W.G., Rubin, D.B., 1973. Controlling bias in observational studies: A review. Sankhyā: The Indian Journal of Statistics, Series A 1, 417–446.
Colussi, T., 2017. Social ties in Academia: A friend is a treasure. Review of Economics and Statistics 100 (1), 45–50.
Comola, M., Prina, S., 2014. Do interventions change the network? a panel peer-effect model accounting for endogenous network changes. IZA Discussion Paper 8641.
Coppock, A., Sircar, N., 2013. Design of field experiments under unknown interference structures. Mimeo.
Cummings, J.N., Kiesler, S., 2005. Collaborative Research Across Disciplinary and Organizational Boundaries. Social Studies of Science 35 (5), 703–722.
Cummings, J.N., Kiesler, S., 2007. Coordination costs and project outcomes in multi-university collaborations. Research Policy 36 (10), 1620–1634.
Dahlander, L., McFarland, D.A., 2013. Ties That Last: Tie Formation and Persistence in Research Collaborations over Time. Administrative Science Quarterly 58 (1), 69–110.
De Stefano, D., Fuccella, V., Vitale, M.P., Zaccarin, S., 2013. The use of different data sources in the analysis of co-authorship networks and scientific performance. Social Networks 35 (3), 370–381.
Del Prete, D., Forastiere, L., Leone Sciabolazza, V., 2019. Causal Inference on Networks under Continuous Treatment Interference: an application to trade distortions in agricultural markets, F.R.E.I.T. Working Paper. pp. 1532.
Ductor, L., 2015. Does Co-authorship Lead to Higher Academic Productivity? Oxford Bulletin of Economics and Statistics 77 (3), 385–407.
Eckles, D., Karrer, B., Ugander, J., 2017. Design and analysis of experiments in networks: Reducing bias from interference. Journal of Causal Inference 5 (1), 20150021.
Fafchamps, M., 2015. Causal Effects in Social Networks. Revue Économique 66 (4), 657–686.
Fafchamps, M., Leij, M.J., Goyal, S., 2010. Matching and network effects. Journal of the European Economic Association 8 (1), 203–231.
Falk-Krzesinski, H.J., Contractor, N., Fiore, S.M., Hall, K.L., Kane, C., Keyton, J., et al., 2011. Mapping a research agenda for the science of team science. Research Evaluation 20 (2), 145–158.
Forastiere, L., Airoldi, E.M., Mealli, F., 2017. Identification and estimation of treatment and interference effects in observational studies on networks. ArXiv Preprint ArXiv:1609.06245.
Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Milojević, S., et al., 2018. Science of science. Science 359 (6379), eaao0185.
Freeman, L.C., 1979. Centrality in social networks conceptual clarification. Social Networks 1 (3), 215–239.
Gesell, S.B., Barkin, S.L., Valente, T.W., 2013. Social Network Diagnostics: A Tool for Monitoring Group Interventions. Implementation Science 8 (1), 116.
Gesell, S.B., Barkin, S.L., Sommer, E.C., Thompson, J.R., Valente, T.W., 2016. Increases in Network Ties Are Associated With Increased Cohesion Among Intervention Participants. Health Education & Behavior 43 (2), 208–216.
Greene, M., 2007. The demise of the lone author. Nature 450 (7173), 1165.
Hackett, E.J., 2005. Essential Tensions: Identity, Control, and Risk in Research. Social Studies of Science 35 (5), 787–826.
Hamermesh, D.S., Oster, S.M., 2002. Tools or toys? The impact of high technology on scholarly productivity. Economic Inquiry 40 (4), 539–555.
Hollis, A., 2001. Co-authorship and the output of academic economists. Labour Economics 8 (4), 503–530.
Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47 (1), 5–86.
Imbens, G.W., Rubin, D.B., 2015. Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.
Jacobs, J.A., Frickel, S., 2009. Interdisciplinarity: A Critical Assessment. Annual Review of Sociology 35 (1), 43–65.
Jacob, B.A., Lefgren, L., 2011. The impact of research grant funding on scientific productivity. Journal of Public Economics 95 (9-10), 1168–1177.
Krackhardt, D., 2001. Viscosity models and the diffusion of controversial innovation. Dynamics of Organizations: Computational Modeling and Organization Theories 22 (2), 268–273.
Latkin, C.A., Knowlton, A.R., 2015. Social Network Assessments and Interventions for Health Behavior Change: A Critical Review. Behavioral Medicine 41 (3), 90–97.
Larivière, V., Haustein, S., Börner, K., 2015. Long-Distance Interdisciplinarity Leads to Higher Scientific Impact. Plos One 10 (3), e0122565.
Leahey, E., 2016. From Sole Investigator to Team Scientist: Trends in the Practice and Study of Research Collaboration. Annual Review of Sociology 42 (1), 81–100.
Leahey, E., Beckman, C.M., Stanko, T.L., 2017. Prominent but Less Productive: The Impact of Interdisciplinarity on Scientists' Research. Administrative Science Quarterly 62 (1), 105–139.
Lee, S., Bozeman, B., 2005. The impact of research collaboration on scientific productivity. Social Studies of Science 35 (5), 673–702.
Leone Sciabolazza, V., Vacca, R., Kennelly Okraku, T., McCarty, C., 2017. Detecting and analyzing research communities in longitudinal scientific networks. PLoS One 12 (8), e0182516.
Long, J.S., Fox, M.F., 1995. Scientific Careers: Universalism and Particularism. Annual Review of Sociology 21 (1), 45–71.
Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60 (3), 531–542.
McCarty, C., Jawitz, J.W., Hopkins, A., Goldman, A., 2013. Predicting author h-index using characteristics of the co-author network. Scientometrics 96 (2), 467–483.
Medoff, M.H., 2003. Collaboration and the quality of economics research. Labour Economics 10 (5), 597–608.
National Institute of Health, 2007. NIH launches interdisciplinary research consortia. Retrieved from. https://www.nih.gov/news-events/news-releases/nih-launches-interdisciplinary-research-consortia.
Paluck, E.L., Shepherd, H., Aronow, P.M., 2016. Changing Climates of Conflict: A Social Network Experiment in 56 Schools. Proceedings of the National Academy of Sciences 113 (3), 566–571.
Rawlings, C.M., McFarland, D.A., 2011. Influence flows in the academy: Using affiliation networks to assess peer effects among researchers. Social Science Research 40 (3), 1001–1017.
Rubin, D.B., 1973. The use of matched sampling and regression adjustment to remove bias in observational studies. Biometrics 29 (1), 185–203.
Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and non-randomized studies. Journal of Educational Psychology 66 (5), 688.
Sá, C.M., 2008. 'Interdisciplinary strategies' in U.S. research universities. Higher Education 55 (5), 537–552.
Shi, X., Adamic, L.A., Tseng, B.L., Clarkson, G.S., 2009. The Impact of Boundary Spanning Scholarly Publications and Patents. PloS One 4 (8), e6547.
Snijders, T.A.B., 2011. Statistical Models for Social Networks. Annual Review of Sociology 37 (1), 131–153.
Sobel, M.E., 2006. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. Journal of the American Statistical Association

101 (476), 1398–1407.
Ugander, J., Karrer, B., Backstrom, L., Kleinberg, J., 2013. Graph cluster randomization: Network exposure to multiple universes. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 329–337.
Vacca, R., McCarty, C., Conlon, M., Nelson, D.R., 2015. Designing a CTSA-Based Social Network Intervention to Foster Cross-Disciplinary Team Science. Clinical and Translational Science 8 (4), 281–289.
Valente, T.W., 2012. Network interventions. Science 337 (6090), 49–53.
Valente, T.W., 2017. Putting the network in network interventions. Proceedings of the National Academy of Sciences 114 (36), 9500–9501.
Valente, T.W., Fujimoto, K., 2010. Bridging: Locating Critical Connectors in a Network Thomas. Social Networks 32 (3), 212–2020.
Valente, T.W., Pumpuang, P., 2007. Identifying Opinion Leaders to Promote Behavior Change. Health Education & Behavior 34 (6), 881–896.
Walker, D., Muchnik, L., 2014. Design of randomized experiments in networks. Proceedings of the IEEE 102 (12), 1940–1951.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of “small-world” networks. Nature 393 (6684), 440–442.
White, D.R., Harary, F., 2001. The Cohesiveness of Blocks In Social Networks: Node Connectivity and Conditional Density. Sociological Methodology 31 (1), 305–359.
Wuchty, S., Jones, B.F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. Science 316 (5827), 1036–1039.
Yegros-Yegros, A., Rafols, I., D’Este, P., 2015. Does Interdisciplinary Research Lead to Higher Citation Impact? The Different Effect of Proximal and Distal Interdisciplinarity. PloS One 10 (8), e0135095.