Institute for New Economic Thinking

Working Papers
ineteconomics.org/workingpapers

# 'Many-Citedness': Citations Measure More Than Just Scientific Impact

Carlo D'Ippoliti*

## Working Paper No. 57

**Revised April 6, 2018**

**ABSTRACT**

Citation indexes are increasingly used to measure the scientific impact of researchers and institutions, though their application is often criticized. Most authors agree that citations do not measure only scientific quality; there is disagreement however on what else they do capture. We study the network of citations of all publications indexed in Web of Science authored or coauthored by Italian tenured academic economists between 2011 and 2015. By estimating the determinants of the probability that any author is cited by any other author in the sample, we find those factors to involve not only similarity in methods and topics but also, significantly, various measures of social community as well as of ideological proximity. We conclude that, at least in the case of economics, citations cannot be interpreted as unbiased proxies of scientific impact, and their use to produce indexes and rankings may require careful rethinking.

**JEL Codes:** A14, I23, D85

**Keywords:** citations, rankings, economics

* Department of Statistical Sciences, Sapienza University of Rome, carlo.dippoliti@uniroma1.it. I wish to thank Valerio Leone Sciabolazza, Daniele Vannella, Giulia Zacchia and Jacopo Temperini for skilful research assistance, and Anna Conte, Orsola Costantini, Thomas Ferguson, Jakob Kapeller, Maria Cristina Marcuzzo, Pia Malaney, Andrea Pisauro, Alessandro Roncaglia, Francesco Sylos Labini, Kevin Young, and the participants at the special session at the *INET 2017 Annual Conference*, for comments and suggestions on previous drafts of this work; all remaining errors are obviously mine. Financial support from the Institute for New Economic Thinking is gratefully acknowledged, INET grant # INO1500040.

## 1. Motivation

In the past decades, the study of science has taken a distinctly quantitative turn due to both the increasing availability of large, rich databases and the development of new techniques of data analysis, with the two trends reinforcing each other (for a recent review, see Fortunato et al., 2018). An especially contested field of inquiry is whether adequate quantitative measures of research quality or impact can be developed on the basis of citations counts. On the one hand, bibliometric indexes based on various ways of counting and aggregating citations are increasingly applied to the evaluation of individuals, journals, departments, universities, and even whole countries (King, 2004). On the other hand, a growing number of researchers, journals, associations and scientific societies distance themselves from these practices (see e.g. the influential joint report from the International Mathematical Union, the International Council of Industrial and Applied Mathematics, and the Institute of Mathematical Statistics: Adler et al., 2008) or at least voice the pressing need for better, unbiased metrics and a more balanced, prudent and self-aware use of them (see e.g. the San Francisco Declaration on Research Assessment, DORA; the Leiden Manifesto for research metrics: Hicks et al., 2005; or more recently the joint declaration by the Académie des Sciences, Leopoldina and Royal Society, 2018).[1]

In this work, we review the literature on the use of citation analysis and its limitations for the creation of indexes aimed at the evaluation of research, and we develop an original application to the case of economics in Italy. We first consider pairs of authors to show that network dynamics, such as being coauthors or sharing research interests, are significant predictors of the probability of citations within each pair of authors. However, other factors are relevant too, including proximate political views. In light of this evidence, we then consider the aggregate number of citations received by each author, finding that various measures of centrality in the networks of authors' affiliations, co-authorship etc., again including political proximity, significantly determine how many citations each author receives in a certain year.

These findings imply that citation-based bibliometric indexes that ignore network dynamics risk being biased, and in general citations counts cannot be considered as unbiased proxies of scientific impact. Furthermore, they cast a shadow on the working of academic economics, in so far as citation patterns and thus visibility and career prospects seem to be significantly affected by an author's political views.

## 2. Two debates on the use and meaning of citation metrics

Citation metrics raise two distinct debates in the literature: empirical, regarding its technical use, and theoretical, regarding its meaning and, more generally, the meaning of scientific impact.

[1] In this review we are only concerned with the use of citation metrics, and not their abuse, such as when journal level indexes are straightforwardly used to measure the impact of single papers, which are then used to infer the impact of individuals, and possibly even aggregated further at the institution, disciplinary or national levels. These practices are considered here as abusive because they apply a certain indicator to measure something different from what it was conceived for.

### *2.1. Empirical challenges*

From an empirical standpoint, the first challenge for citation analysis is that citation counts are easily found to be skewed and biased, in the sense that they correlate with many things beside quality (MacRoberts and MacRoberts, 1996). To mention just some of the known problems (see e.g. Bornmann et al., 2008), at the publication level citations are found to correlate with the number and reputation of the publications' authors, publication age, language, the kind of publication (review articles, editorials, studies using primary data, etc.), the reputation of the journal, the number of pages, and even with title length (Letchford et al., 2015); at the author level, citations depend at least on academic age, field and degree of specialization, and gender (King et al., 2016); and systematic differences are found in the average citations of different disciplines or even fields within the disciplines (Radicchi et al., 2008). Additional sources of bias are, to mention just a few, self-citations, selective and/or implicit citations, the increase in the total number of citations with time, and the fact that several widely used bibliometric indexes, for example the *h* index, are not robust to even trivial changes in the papers or citation counts (Hicks and Melkers, 2012) and/or these indexes themselves correlate with variables that are not related to scientific quality.

Finally, the distribution of citations is problematic as well: not only is it generally highly asymmetric, reducing the significance of mean values, but it also exhibits fat tails, implying both an extraordinary number of papers that are never cited (for the case of economics, see Oswald, 2007) and a considerable number of works that are cited many times more than what many believe would be explained by their intrinsic quality (for a review, see Perc, 2014). These findings are typically associated with the 'Matthew effect', that is the growing polarization whereby works that are already highly cited are more likely to be further cited, enjoying what may be called a cumulative advantage.

#### *2.1.1 The literature on the use of citation indexes*

The problems sketched above have given rise to three streams of literature. One is concerned with the statistical analysis of citation distributions and the selection of the best way to represent and describe them (for the case of economics, see Tol, 2013). For example, Wang et al. (2013) propose a model to collapse the citation histories of papers into a single curve, and find that while the citations of small-impact papers can be described by a lognormal model, the accumulation of citations to high-impact papers is best described by "preferential attachment", that is the name most commonly used in a network analysis setting for the concept of cumulative advantage, whereby the growth of citations to a paper is proportional to its accumulated past citations.

A second stream of literature tries to develop better measures, refined indexes and rankings, and the use of new or different data. For example, Radicchi et al. (2008) propose the use of standardized citation counts to level out the systematic differences across disciplines, while Corsi et al. (2011) propose a similar procedure by research field within economics.[2] Harnessing the new possibilities offered by online scientific publication and communication, the "Altmetric" approach proposes the adoption of several distinct metrics and qualitative data that

[2] Indeed, for example in the Italian context the evaluation of single researchers' citations for the aims of recruitment and promotions is standardized by within-discipline research field, but this only applies to the natural sciences.

are complementary to traditional, citation-based metrics such as article downloads, abstract views, etc. (Sud and Thelwall, 2014).

Finally, the third stream of literature focuses on the consequences of citation-based measures. From a narrow perspective, these consequences include changes in the patterns and distribution of citations themselves, arising from scientists' reaction to the chosen metrics of evaluation. A long-run increase in citations and self-citations is observed in all disciplines in which bibliometrics has become the dominant tool of research evaluation (King et al., 2016). Various studies link the pressure to publish in impact journals with growing malpractice and unethical behavior (e.g. Brembs et al., 2013) and there are strong indications that the incentives structure created by the simplistic use of bibliometrics as an evaluation tool may induce undesired outcomes (Edwards and Roy, 2017). Among the latter, it is notable that citations themselves may have lost their role as indicators of quality or even scholarly impact, as implied by the so-called Goodhart's law.[3] For example, an anonymous survey on 426 economists found that 52% of respondents failed to read the content of works they cited, and 20% deliberately refrained from citing works published on low ranked journals (Necker, 2014). While these behaviors may be deemed scientifically unethical, we are not specifically concerned here with illicit behavior, such as peer review fraud, "citation rings", predatory journals, fake research, or citation coercion, even though some of these may be especially problematic for economics (see e.g. the annex in Wilhite and Fong, 2012) and have been known to be growing for some time now (for the case of management studies, see Macdonald and Kam, 2010).

From a wider perspective, the reflexivity between citation metrics and scientists' behavior is not limited to citation practices: it extends to research activities too (Hicks and Potter, 1991). In the survey on economists by Necker (2014), 67% of respondents declared they choose a topic of research on the basis of the perceived prospects for publication of the expected results – a behavior that for example Pfeiffer and Hoffmann (2007) find in the literature on genetics as well. In general, citation measures are found to discourage multidisciplinary and 'frontier' studies (see e.g. Rafols et al., 2012). In economics, the use of citation metrics that are not normalized is found to imply strong value judgments on the relative merits of various methods and fields of research, e.g. to the detriment of the history of economic thought in favor of econometrics, as well as to strongly discourage the study of peripheral economies in favor of the investigation of mainly the US one (Corsi et al., 2011). More generally, the use of citation-based metrics for evaluation, hiring and promotion purposes risks seriously reducing the freedom of researchers, especially younger ones, to choose what topics and methods to study due to their intrinsic interest, with potentially disruptive consequences on the future development of their discipline.

As a consequence of these findings, a debate has ensued about the actual role of bibliometrics in causing these outcomes. A recent example revolved around the Italian case, in which the evaluation of university departments and research centers relied on the mechanical application of bibliometrics to an unusual degree, while mandating that a sample of works be evaluated through peer review too, for the aim of comparison. Using undisclosed data, members of the evaluation panel published a study in which they claim no significant differences between citation metrics and peer review emerge in the predicted assessment of a publication (Bertocchi et al., 2015). However, the statistical soundness of their conclusions was disputed by Baccini

[3] In the economic and management literatures, Goodhart's law denotes the tendency for a chosen measure to lose its significance once it is adopted as a target.

and De Nicolao (2016), to which a comment by Bertocchi et al. and a rejoinder by Baccini and De Nicolao followed in the same journal. This debate resembles a larger one, on the general agreement between evaluations based on bibliometrics and on peer-review, in which empirical studies usually find weak correlation between the scores assigned by the two methods (Moed, 2005).

### *2.2. The debate on the meaning of citations*

Beside the empirical challenges listed in the previous section, a second – possibly more destructive – challenge for citation analysis is theoretical. This concerns the interpretation of its results. Most authors agree that references and citations do not measure only scientific 'merit', 'quality' or 'relevance'; there is disagreement however on what else do they capture, and whether this multidimensionality prevents the use of citation analysis as a measure of intellectual accomplishment (Hicks and Melkers, 2012). Evidently, the answer to this question depends on one's definition of scientific 'impact'. The most common use of the term is usually that proposed by Martin and Irvine (1983), according to whom referencing arises from several possible motivations, and so citation counts measure both intellectual and social influence at the same time (hence the use of the term *impact* rather than *quality*).[4]

One main source of disagreement on the meaning of citation counts lies in the double nature of references. In their textual context, they have a rhetorical function: they convey information about contents, as well as legitimacy through the appeal to another scientist's work. However, by being part of a list of references, which then become citations, they have the independent function both of contextualizing an author's contribution and of providing prestige to others and useful 'points' in the 'citations game'. Moreover, as Camacho-Miñano and Núñez-Nickel (2009) argue, a second crucial issue is the existence of a maximum number of references that can be included in a paper. For this reason, they propose to distinguish the process of references selection into two steps (after a preliminary one, of excluding non-citable works): first, a researcher collects all studies that in an objective way may be considered as relevant to her own work; then, from this pool she picks those that she will actually cite, necessarily choosing in a discretionary way. Thus, even assuming that an objective choice at some stages is possible, the final outcome is nonetheless subjective.

Empirical investigations of individual citers' motives usually reflect this ambivalence in the role of citations, finding that both intellectual and "ceremonial" motivations drive citation behavior (Bornmann and Daniel, 2008). In economics, a qualitative study on two samples of 25 and 15 agricultural economists highlighted the relevance of power relations within the discipline (White and Wang, 1997).[5]

---

[4] Moed (2005) distinguishes five non-mutually exclusive approaches to the study of citations: a physical, sociological, psychological, historical and an information-scientific viewpoint. In his taxonomy, the physical approach argues that citation indicators should be interrelated by simple quantitative laws in order to reach an understanding of their phenomenology. The sociological approach focuses on the motivations of scientists, and their interactions with colleagues and policy makers; connected to this approach, the psychological one inquires into individual citers' motivations. The historical studies focus on the use of bibliometric data to ascertain the past development of scholarly ideas and their principal contributors, while the information-scientific approach studies the concept and diffusion of information within the scientific literature.

[5] They quote researchers reporting about citations in their published works: "we didn't want to be told we had neglected to cite certain people. So there are people in here, for example, X is one of these people we anticipated being a referee" (p. 145), or "[i]n economics there are all different kinds of levels of journals, and the theoretical level that we were aiming at is most closely matched by the *Journal of Economic Theory*, *Review of Economic Studies*, and *Econometrica*. The paper that we actually wrote was ultimately submitted to *Econometrica*. So, when we picked out references, we tried to stay in that group. It is a little bit of gamesmanship in a way, to be citing the right people" (quoted in White and Wang, 1997, p. 136).

It may be that disciplines differ in the relative frequency of intellectual and ceremonial citations. For example, according to Krampen et al. (2007) perfunctory citations, that is citations not required for the understanding of the text or to sustain the argument, are more frequent in psychology than in the social sciences or physics.

However, in light of the reflexivity of citation analysis, these findings are not likely to remain the same as bibliometrics develops and becomes increasingly institutionalized. An indication may be provided by Case and Miller (2011), who contrast citation habits in a sample of 112 scholars among whom 63 self-described 'bibliometricians'. They find that bibliometricians are significantly more likely than the others to reference a work because it is authored by a "recognized authority in the field" or it was "published in a prestigious journal" (p. 426). It is thus possible that as citation analysis spreads across disciplines (from the natural to the social sciences to the humanities) more scientists will become similar to bibliometricians in their awareness that references are countable citations too, and the observed differences in the relative frequency of intellectual and perfunctory citations across disciplines will level out as well. From this point of view, economics currently stands in a middle ground between the natural and the social sciences, with bibliometrics increasingly applied to the evaluation of economic research amid strong opposition to it in many countries.

However, distinguishing which citations are merely perfunctory and which instead serve a scientific purpose may not always be possible (Davis, 2009). For example, Moed (2005, p. 217) proposes a model of citation behavior that reflects both intellectual and social considerations. He assumes that science is organized into research groups, which produce two sorts of papers: 'bricks' and 'flags'. The vast majority of papers, the bricks, contain "normal" contributions, whereas flag papers present either significant new results or overviews of the group's research program: either way flags become symbols of the whole group's progress, approach and range of studies, even beyond the content of the specific paper and/or the identity of its authors. As a consequence, flag papers end up attracting significantly more citations than would be warranted by their intellectual contribution only, because researchers would typically cite one or more flag papers to imply or acknowledge the whole group's research. In this scheme, citations to flag papers do not reflect their intrinsic 'quality' only, but they would be not just perfunctory either.

Two crucial issues emerge within this literature. First, scientists' rationality and agency must be taken into account. Thus, citations are not meaningless, even though they do not reflect scientific quality in any straightforward way (Hicks and Melkers, 2012).
At the individual level, this issue connects to the longstanding debate over scientists' motivation and (dis)interestedness (Davis, 2009). At the aggregate level, scientists' behavior may produce various kinds of outcome depending on the context.

Moed (2005) argues that citations reflect both intellectual and social prestige:

> "In any field there are leading groups active at the forefront of scientific development. Their leading position is both cognitively and socially anchored. Cognitively, their important contributions tend to be highlighted in a state-of-the-art of a field. But *to the extent that the science system functions well* in stimulating and warranting scientific quality, leading groups, and particularly their senior researchers, tend at the same time to acquire powerful social positions." (p. 219, italics added).

Thanks to the efficiency of science, there is thus an indirect channel whereby even perfunctory citations and references inserted in a text for social reasons, ultimately reflect the outstanding scientific achievement of the cited person. Reliance on this indirect channel, though, rests on the assumption that science is in fact organized in a way that rewards scientific achievement by conferring prominent social positions to the individuals who contribute most to the scientific development of their field, and to them only. Ultimately, the opinion on what citations measure depends on one's views on the working of the science industry.

These considerations are linked to the second issue discussed in this literature, that citation analysis is not usually aimed at capturing individuals' motives, but rather aggregate trends. Thus, it is necessary to distinguish two sorts of 'error' when equating citations with scientific impact (if we recognize that the science industry is not perfectly efficient). A first sort of error may reflect individual citers' idiosyncrasies, and may thus likely be reduced or even eliminated by the simultaneous consideration of large datasets in which several people's idiosyncrasies will cancel out. A second sort of error reflects instead systematic trends, that is, behavior correlated across individuals, and should thus be recognized as "bias" in the sense that larger datasets are not less affected by it than a single publication. Thus, the degree to which citations approximate scientific quality depends on how much the reasons for citing a paper, other than its scientific contribution, are correlated across citers.

Indeed, even if biases may be empirically detected, it does not follow that unbiased measures of scientific impact can be easily created. The concept of bias, that is, of measuring something different from what was sought, is itself part of the definition of what should really be measured. Unless one defines scientific quality one cannot know what is a bias, and vice versa, by labeling something as a bias we implicitly define at least what scientific quality is not. As a consequence, citation measures as well as the proposed corrections of their biases inextricably reflect value judgments.

In this work we deal with both issues described here, providing evidence on some systematic patterns of scientists' individual citing behavior as well as some of their consequences at the aggregate level. We consider all economists based in Italy in the period 2011-2016, and extend the dataset to their main social and scientific connections, as explained in section 4. Even though we stop short of fully defining what constitutes scientific quality in economics our work has normative connections, for it labels some observed trends as sources of bias; that is, we identify some aspects that we deem at any rate external even to an implicit understanding of scientific quality. Indeed, we exploit network connections between economists to highlight some forms of bias in the form of citations that are unrelated to scientific quality even from an indirect point of view in the sense of Moed (2005), according to which perfunctory citations may reflect the cited person's scientific merit as evidenced by her social position. Furthermore, we document that these biases do not cancel out in the aggregate. Thus, the system of economic research, in Italy and possibly elsewhere, cannot be assumed to "function well" and the connection between citations and scientific impact should be regarded as loose at best.

## 3. The literature on citations networks

Network analysis is increasingly used by sociologists and economists to represent and analyze

social interactions, including scholarly interactions among economists, on which we focus here (Goyal, 2016; Jackson et al., 2017). The bulk of these works consider networks of co-authorship among economists (Goyal et al., 2006; Fafchamps et al., 2010; Cainelli et al., 2012; Ductor et al., 2014; Besancenot et al., 2016; Molina et al., 2016) or in the finance field (Georg and Rose, 2016a, 2016b); some also consider their potential impact on other variables, such as researchers' productivity.

In their first contribution, Goyal et al. (2006) show that, despite the growth in active researchers, economics exhibits the characteristics of a "small world", primarily a short average distance between any two members of the co-authorship network. As the "giant component" of connected authors grew between the 1970s and the 1990s, and isolated authors shrank in parallel, their main finding is that economics appears to be structured into small communities, connected by "interlinked stars". By the latter they mean an economist who writes with many other economists, most of whom have few coauthors and generally do not write with each other. These stars effectively bridge different communities, which gives them high 'network centrality'. Indeed, works in this field traditionally distinguish at least four measures of centrality, denoting by "node" the components of a network and by "links" the connections between them. For each node (author), *degree centrality* is defined as the number of connections (co-authorships) that the author has with any other author; *PageRank centrality* is a recursive notion, in which the links that a node has (her co-authorships) are weighted differently according to the centrality of the nodes with which these links are established (the centrality of these coauthors); *closeness centrality* is the average of the shortest distance between the author and all other authors; and the *betweenness centrality* of a node is defined as the proportion of the shortest paths between all pairs of authors in the network that pass through that node. Goyal et al.'s (2006) interlinked stars are economists with extraordinary high degree centrality in the co-authorship network.

Subsequently, Fafchamps et al. (2010) analyzed nearly the same phenomenon from a different point of view, finding that in the co-authorship network of economics the distance between two authors is inversely correlated with their probability of establishing a new connection (publishing a work together for the first time).

Cainelli et al. (2012) try to understand the impact of the heterogeneity in the economists' propensity to write with coauthors on their productivity. They consider all Italian economists in 2006 and estimate their output (as measured by the number of publications in EconLit) as a function of individual characteristics as well as some "relational" variables, namely their propensity to cooperate and the international reach of the individual's co-authorship network. They find that co-authorship is a significant determinant of scientific productivity, but it should be noted that their measure of the 'propensity to cooperate' is not directly estimated through network analysis, but rather it is obtained through instrumental variables. Specifically, they consider an author's attitude to write book chapters as a measure of her propensity to cooperate because it would show both a connection between the author and the book editors, and because it may imply some selflessness on the part of the author, to the extent that writing book chapters reduces the time available for writing journal articles, which are the only publications considered in most citation and bibliometric indexes (for the case of economics, see Corsi et al., 2011).

More recent articles conceptualized and measured relational variables as network connections,

in order then to assess the impact of these connections on scientific productivity. Considering all journal articles indexed in EconLit between 1970 and 1999, Ductor et al. (2014) find that incorporating information about the co-authorship network leads to a statistically significant though quantitatively modest improvement in the accuracy of forecasts of economists' output. In a sample of all Spanish economists and their coauthors between 2002 and 2014, Molina et al. (2016) too find that network centrality and scientific productivity are correlated. Similarly, a positive impact of co-authorship on productivity is found in a sample of French economists by Besancenot et al. (2016), who further control for the possibility of assortative matching, i.e. the hypothesis that a highly productive author should more frequently meet authors willing to collaborate with her, and should thus have more co-authored papers than authors with low productivity (which would introduce endogeneity issues in the estimate).

In their analysis of productivity, Besancenot et al. (2016) partly consider citations as a dependent variable, in so far as they adopt the *H* and *G* indexes as measures of output rather than the sheer number of published items. But in general, to our knowledge citations networks in economics have not been empirically studied yet. They were rather analyzed by a series of works that focused on the natural sciences or occasionally on some other social science (Baldi, 1998; Mählck and Persson, 2000; White et al., 2004; Johnson and Oppenheim, 2007; Yan and Ding, 2009; Wallace et al., 2012; Li et al., 2013; Uddin et al., 2013; Abbasi et al., 2014).

Considering a subfield of astrophysics, Baldi (1998) runs a dyadic logistic model considering all possible pairs of authors. He finds that the content of the cited article and its "quality" are predictors of the citations that an author receives from another one (he measures quality by the total number of citations received by the cited author, excluding self-citations and those by the citing author). Instead, an author's position within the stratification structure of science and some relational variables between the two authors do not appear to significantly improve the fit of his model. Baldi interprets this evidence as implying that citations reflect payment of intellectual debt and not other social dynamics. Similarly, White et al. (2004) analyze public and private communications within an interdisciplinary group of researchers working on human development, finding that shared content between two documents is a better predictor of citations than friendship between their authors, even though, considering the intensive margin, the authors in their sample who exchanged citations (the "interciters") appear to cite each other more than the others.

In contrast to these findings, considering two departments of biology in Sweden Mählck and Persson (2000) find considerable overlap between the co-authorship and the citation network. Similarly, considering the extended citation networks of three information scientists, Johnson and Oppenheim (2007) find a positive correlation between social closeness (measured through individual questionnaires) and citation counts, despite their finding that information scientists cite widely outside their immediate social connections too.

These conflicting results could be explained by variations across disciplines. As Wallace et al. (2012) note, scientific fields vary considerably both in co-authorship and in citing practices, for example depending on the mean number of coauthors per paper (which in some disciplines can reach several dozens), or the fact that the field is more or less fragmented into separate communities working on different topics. Thus, citations between (previous) co-authors are generally lower in the social sciences than in the natural sciences because co-authorship is less frequent and each author publishes considerably less articles: "as a consequence, researchers

have less co-authors in their social network to choose from" (p. 3) when citing.

Finally, even if social connections did not exert a significant impact on citations between two specific authors, it could be that an author's position in the network has an impact on her total citations. Li et al. (2013) and Abbasi et al. (2014) frame this intuition in terms of the author's social capital, which provides benefits beyond the individual ties she has. From this perspective, similarly to the studies on the impact in terms of productivity, some authors try to estimate the impact of an author's position within the co-authorship network on her citation counts. Specifically, Yan and Ding (2009), Li et al. (2013), and Uddin et al. (2013) find that various measures of centrality (closeness centrality, betweenness centrality, degree centrality, and PageRank) are significantly correlated with citation counts. Abbasi et al. (2014) obtain the same result for both citation counts and authors' *h* index. According to Li et al. (2013), in a network of 137 information systems scholars, *betweenness centrality* plays the most important role because it would allow authors to exploit "non-redundant resources" in terms of social capital within their co-authorship network. This idea is similar to Goyal et al.'s (2006) concept of "star economists", even though the latter notion was developed to denote the economists exhibiting extraordinary high *degree centrality*.

Partly, these network dynamics may be confused with other trends when analyzing citation counts at the paper or author level without a network framework. For example, researchers with longer publishing tenure tend to have higher degree centrality, and it may be difficult to distinguish which factor is actually exerting a positive effect on the number of citations received by an author. To tackle this sort of issue, in the next two sections we develop an analysis of the impact of an author's network connections on her citation counts looking both at dyadic citations (i.e. from a specific citing author to a specific cited author) and in the aggregate.

## 4. Data and Methodology

Data for this study were gathered by querying: (a) the Web of Science Core Collection (owned by Clarivate Analytics, formerly Thomson Reuters), a large subscription-based bibliographic database, which was selected because it has complete and consistently formatted citation information for its entries; (b) the Italian Ministry for Education, Universities and Research database of university staff; and (c) an ad hoc archive of all Italian economists' contributions on Italy's major print and online media outlets.

We considered all 948 tenured academic economists affiliated in an Italian university in at least one year between 2011 and 2015.[6] In the Italian system, these are either professore ordinario (roughly corresponding to a full professor), professore associato (associate professor or senior lecturer) or ricercatore universitario (lecturer or assistant professor) classified by the Italian Ministry for Education, University and Research (MIUR) as working in the field of economics (formally referred to as "political economy, SECS-P/01").[7] We searched for all publications

[6] The full list of academics with tenure in an Italian university, by year, rank and discipline, was obtained from the Italian Ministry for Education, Universities and Research, at: http://cercauniversita.cineca.it/php5/docenti/cerca.php

[7] In the Italian context, some scholars that would be considered economists in an international context may in fact be classified as working in other fields, such as econometrics, economic statistics, public policy, public finance or others. While many of them enter in the analysis as some economist's connection, it was not possible to include all of them because in these fields

indexed in Web of Science in the relevant period, by an author or coauthor with the surname and initial of the first name corresponding to each economist in the above defined group. To limit the possibility of collecting publications by other, same-named authors, we limited the search to all works indexed in the subject category "economics," and in some cases we checked for possible homonyms by looking at the authors' affiliations and by comparing the results with information from other sources (personal and institutional websites, publicly available CVs, etc.).[8]
From all subsequent analyses we excluded authors for the years prior to their first publication indexed in Web of Science, because citations from/to their works would be missing for reasons (non-inclusion in the database) different from those related to authors who published at least one paper and yet were not cited. For each author as a citer, we further exclude from the analysis the years in which she did not publish at least one paper indexed in Web of Science because, by definition, in those years she would have no opportunity to cite anybody.

For 775 tenured Italian economists we were able to find at least one publication indexed in Web of Science in the subject category "economics" in the relevant period. Their papers were cited 1969 times, of which 621 citations were made by at least another Italian tenured professor of economics (since papers are often co-authored a same citation can count as a link between more than two authors). In total 142 economists cited and 151 were cited by another economist in the sample at least once in the period considered.

Beside citations and co-authorship we collected information on each paper's title, journal, abstract and references, and on the authors' self-declared institutional affiliations. The primary affiliation (that in which they have tenure) was obtained from MIUR's website, but authors may have multiple affiliations, for example if they are affiliated with research centers (such as CEPR, NBER, etc.) and because a non-negligible number of tenured faculty employed in Italian public universities more or less continuously work as adjunct faculty in private or other public universities too (D'Ippoliti and Zacchia, 2017).[9]

A prime reason for citing a scientific paper is if it deals with a topic similar, connected, or complementary to that on which someone is working. With the aim of measuring the distance between the topics and methods adopted by two authors, for each author in each year we aggregated the titles and abstracts of all the papers she published until that year. These texts were used to compute the cosine similarity of each pair of authors' scientific production. In the information science literature the cosine similarity of a text is a widely used measure of the overlapping of two or more texts, based on the normalized share of common used terms (for an application in a network setting see Fafchamps et al., 2010). However, due to the brevity and

---

there are also several non-economists such as statisticians, political scientists, law scholars, etc. and it is not possible to define the economists in an objective way.

[8] It was necessary to search the full surname and full first name, due to their sharing the same surname and initial of the first name with one or more other researchers indexed in the database for the following authors: Chapman Sheila, Conte Andrea, Gallegati Mauro, Gallegati Marco, Lanza Giuseppe, Lombardi Mauro, Moro Andrea, Moro Alessio, Motta Gregorio, Panico Carlo, Panico Claudio, and Patalano Rosario. Furthermore, author Paolo Giordani had to be excluded from the analysis due to the impossibility to distinguish some of his works from those of a same-named statistician, similarly employed as a tenured academic in an Italian university, who may have published works indexed in the economics subject category of Web of Science.

[9] Affiliations were not always reported, and in several cases were reported with slightly different names on different publications. For this reason, we applied both textual analysis and manual checking to match the various names of the same institutions. Moreover, since Italian university staff, especially tenured faculty, exhibit very low geographical mobility, for each author after the first published paper we assume that the affiliation(s) subsequently remained the same until the publication of their subsequent paper, at which point the affiliation may change or remain the same.

specificity of metadata such as economics journal abstracts, by this measure papers on fairly close topics in our sample may have appeared as sharing nothing in their textual description. For this reason, we preliminarily devised a list of words that, to our aims, may be considered as synonymous and replaced them in each author's texts. This list was developed into two steps. First, all keywords used as official descriptors of a same JEL code were flagged as synonymous,[10] (considering compound keywords such as "monetary policy" or "economic history" as a single term). Second, a list of the 400 most commonly used words in the sample was manually processed with the aim of aggregating synonymous terms. A list of all aggregated terms is reported in Appendix 1.

Finally, for all tenured professors of economics in an Italian university, we collected the number, dates, titles and source of all the articles, interviews, commentaries, letters, and op-eds they wrote on the 68 main national newspapers and weekly and monthly magazines, as well as the articles and posts written on 10 widely read economics blogs and online magazines (for a full list, see Appendix 1). Even by international standards, Italy is considered to exhibit a media system highly polarized along political lines (Mancini, 2013) and, as discussed for example by Helgadóttir (2016), Italian economists have played a relevant role in shaping policy making at a European level in the period considered in our analysis, inter alia by their contributions to the national and international public discourse on the press. Therefore, for economists, contributing to a certain outlet may denote political or ideological proximity to its editorship and/or readership, and writing on the same outlet may denote a political or ideological connection between two economists. According to some authors, joint institutional affiliations too may denote ideological proximity (Helgadóttir, 2016; Zacchia, 2016).

In each year, the networks formed along these various dimensions are not very dense, with the exception of the network of topics similarity, which every year exhibits density around 0.18. However, since these networks do not exhibit a strong overlap, considering more than one simultaneously allows us to create a denser network, as exemplified in figure 1 (for the other years, see Appendix 2). In year 2015, the network of coauthorships (shown in panel A) has 287 nodes connected into 18 components; the network of common institutional affiliations (panel B) has 184 nodes and 68 connected components; the network of common media outlets (panel C) has 86 nodes connected into 1 main component; and the network formed by the union of the previous three (panel D) has 353 nodes connected into 7 connected components.

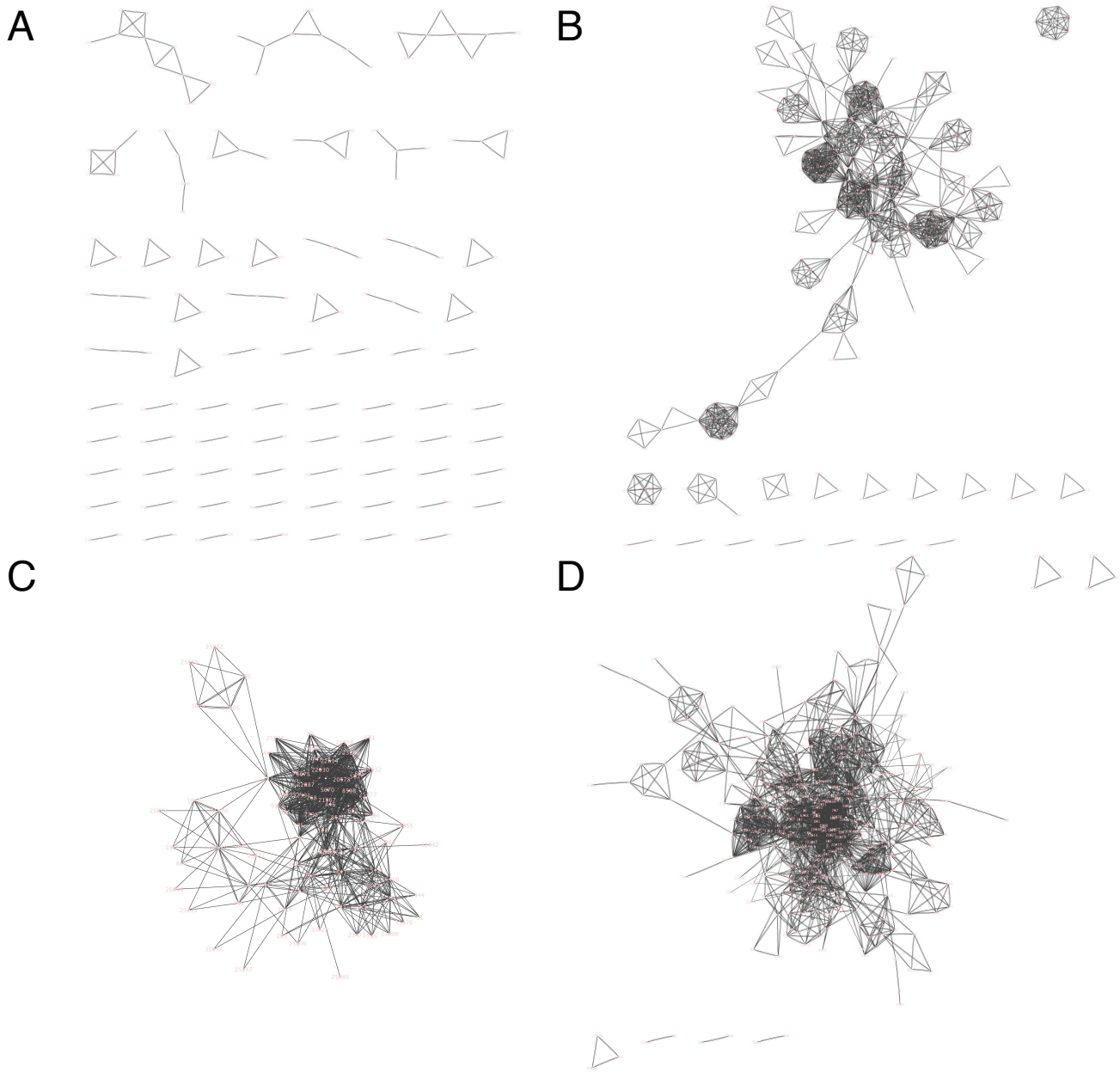
After excluding authors as potential citers for each year in which they do not publish at least one publication, and as potentially cited authors for all years before their first recorded publication, we observe 1,148,348 possible dyadic interactions between 607 potential citing authors and 775 potential cited authors.

With these data we first estimate a model of network formation at the individual level, looking at the probability that an author $i$ in the sample cites another author $j$ in the sample in a certain year, $t$, denoted by $\Pr(C_{ij,t})$. We exclude self-citations from the analysis, thus $i \neq j$, and we assume dyadic independence, i.e. the fact that $i$ cites $j$ does not affect in an unobserved way the probability that $k$ cites $j$, for all $i$'s, $j$'s and $k$'s in the sample.

[10] *Journal of Economic Literature* (JEL) classification system codes are alphanumeric codes, periodically updated by the American Economic Association, used to identify subfields within economics. A full list is available at https://www.aeaweb.org/econlit/jelCodes.php?view=jel. For the keywords repeated across several JEL codes, we attributed the synonymous term(s) to the JEL code of their first occurrence.

Such model could be conceptualized in two different ways. On the one hand, the probability that *i* cites *j* in a certain year could be correlated with the probability that she cites the same author in another year. In other words, there may be dyadic unobserved characteristics that relate *i* and *j*, beside those explicitly accounted for in the model. To control for such possible correlation in the residuals, we estimate a conditional fixed-effects logistic regression model, taking each pair (dyad) as a unit of analysis.

**Figure 1.** The connected components in the networks of coauthorship (A), common institutional affiliations (B), common media outlets (C), and in the union of all three (D), for year 2015.



On the other hand, it may well be that citations from one author to any other author in the sample, and/or from any author to a specific one, are correlated over time. This could happen if authors have unobserved individual characteristics that influence how many citations they make and/or receive (Graham, 2017). Therefore, in a second specification we estimate a pooled logistic regression model, estimating 2-way cluster-robust standard errors, allowing for correlation in the residuals by citer and separately by cited author.

To account for these two non-nested sets of clusters, we use a Huber-White sandwich estimator of the coefficients' standard errors based on a generalization of the cluster-robust variance matrix. Specifically, following Cameron et al. (2011) robust standard errors are obtained by adding the separately estimated variance-covariance matrices of the estimators obtained with standard errors clustered on the first set of clusters (citing authors), $\hat{V}_1[\hat{\beta}]$, and on the second set (cited authors), $\hat{V}_2[\hat{\beta}]$, and subtracting the variance-covariance matrix of the estimators with errors clustered on the intersection of the two sets (pairs of authors), $\hat{V}_{1\cap 2}[\hat{\beta}]$. The estimated variance matrix is thus $\hat{V}[\hat{\beta}] = \hat{V}_1[\hat{\beta}] + \hat{V}_2[\hat{\beta}] - \hat{V}_{1\cap 2}[\hat{\beta}]$.
As a robustness check, robust standard errors were obtained with 2,000 block bootstrap repetitions as well (Cameron et al., 2008).

In all specifications, the set of observed dyadic attributes at time $t$, $X_{i,j,t}$, includes (ignoring the $i$ and $j$ lower scripts for simplicity):

- Two measures of proximity between $i$ and $j$, in terms of belonging to a scientific community or research group in the sense recalled in section 2.2: the number of jointly written papers ($P_t$), and the number of common institutional affiliations ($A_t$);
- Two measures of similarity of research topics and/or methods: the number of journals in which both $i$ and $j$ published at least one article ($J_t$), and the cosine similarity of the metadata of all papers written by $i$ and $j$ up to time $t$ ($S_t$);
- and a measure of political or ideological proximity between $i$ and $j$: the number of newspapers, magazines and blogs in which both $i$ and $j$ gave at least one interview or wrote at least one article, commentary or post ($POL_t$).

We further include as control variables time (in a linear trend), the number of publications by $j$ (cited author) indexed in Web of Science up to time $t$ and therefore citable, and the number of papers written by $i$ (citing author) and indexed in Web of Science in $t$, which may capture $i$'s opportunities to cite $j$. Descriptive statistics for all variables considered are reported in Appendix 2.

## 5. Results and discussion

Table 1 reports the main results, as well as a robustness check carried out by including lagged explanatory variables. As shown in the table, when controlling for either dyadic or individual clustering the results of the estimates do not qualitatively change.

The use of bootstrapped standard errors, reported in columns 2 and 5, reduces the size of all coefficients' estimated standard errors, denoting higher statistical significance of the coefficients (the results in columns 1 and 4, obtained with the method developed by Cameron et al., 2011, should therefore be regarded as more conservative). When estimating the panel

conditional effects logistic model (columns 3 and 6), the size of the estimated coefficients is generally higher than in the pooled estimates, while their standard errors are often roughly similar. Accordingly, the conditional effects panel model suggests a slightly stronger role for network connections in determining dyadic citations than the pooled model.

Among the control variables, time is positively correlated with the probability of citations, denoting a trend of 'citations inflation' due to both the increase of citations per author over time, and the increase of economics journals indexed in Web of Science (citable items). The number of an author's past publications appears to be correlated with her probability of being cited, though not when controlling for lagged independent variables; and writing more in a certain year generally correlates with the probability to cite another author in the sample.

**Table 1. Network Formation: logistic model of the probability of dyadic citation (odds ratios)**

| | (1) 2-way cluster-robust pooled | (2) Bootstrap SE pooled | (3) Conditional effects panel | (4) 2-way cluster-robust pooled | (5) Bootstrap SE pooled | (6) Conditional effects panel |
|---|---|---|---|---|---|---|
| Connections in the journals network in *t* | 3.293*** | 3.293*** | 3.652*** | 3.242*** | 3.242*** | 4.334*** |
| | [0.353] | [0.233] | [0.349] | [0.499] | [0.340] | [0.714] |
| Conn. in the coauthorship network in *t* | 5.201*** | 5.201*** | 6.439*** | 2.533* | 2.533*** | 4.529*** |
| | [1.495] | [1.078] | [2.182] | [1.219] | [0.754] | [2.224] |
| Connections in the affiliations network in *t* | 1.871*** | 1.871*** | 2.209*** | 0.696 | 0.696 | 1.041 |
| | [0.451] | [0.355] | [0.572] | [0.485] | [0.349] | [0.657] |
| Connections in the topics network up to *t* | 2.921** | 2.921*** | 2.825** | 1.461 | 1.461 | 1.990 |
| | [1.218] | [0.886] | [1.191] | [0.786] | [0.582] | [1.311] |
| Connections in the media network in *t* | 1.045* | 1.045** | 1.052** | 1.081*** | 1.081*** | 1.089*** |
| | [0.0254] | [0.0212] | [0.0237] | [0.0238] | [0.0163] | [0.0327] |
| Conn. in the journals network in *t* - 1 | | | | 2.293*** | 2.293*** | 2.694*** |
| | | | | [0.543] | [0.236] | [0.485] |
| Conn. in the coauthorship network in *t* - 1 | | | | 2.819** | 2.819*** | 5.122*** |
| | | | | [1.300] | [0.741] | [2.180] |
| Conn. in the affiliations network in *t* - 1 | | | | 1.482 | 1.482 | 1.455 |
| | | | | [0.891] | [0.701] | [0.888] |
| Conn. in the topics network up to *t* - 1 | | | | 2.565 | 2.565** | 4.410** |
| | | | | [1.562] | [1.051] | [3.072] |
| Conn. in the media network in *t* - 1 | | | | 0.970 | 0.970 | 0.945 |
| | | | | [0.0613] | [0.0624] | [0.0979] |
| Publications by cited author up to *t* | 1.049*** | 1.049*** | 1.047*** | 1.001 | 1.001 | 1.002 |
| | [0.0142] | [0.00839] | [0.0120] | [0.0254] | [0.0143] | [0.0227] |
| Publications by citing author in *t* | 1.116 | 1.116** | 1.160** | 1.205** | 1.205*** | 1.297*** |
| | [0.0867] | [0.0576] | [0.0703] | [0.0966] | [0.0620] | [0.0852] |
| Year (linear trend) | 1.774*** | 1.774*** | 1.989*** | 1.568*** | 1.568*** | 1.746*** |
| | [0.132] | [0.0711] | [0.119] | [0.192] | [0.110] | [0.177] |
| Ln ($\sigma^2$u) | | | 4.602*** | | | 7.699*** |
| | | | [0.648] | | | [1.317] |
| Constant | 1.28e-05*** | 1.28e-05*** | 7.27e-07*** | 2.10e-05*** | 2.10e-05*** | 2.18e-07*** |
| | [4.93e-06] | [2.51e-06] | [3.32e-07] | [1.22e-05] | [6.76e-06] | [1.66e-07] |
| *Observations* | *1,148,348* | *1,148,348* | *1,148,348* | *494,258* | *494,258* | *494,258* |
| *Number of pairs* | | | *462.613* | | | *257,270* |
| *Standard errors clusters* | *607 citing, 775 cited* | *607 citing, 775 cited* | | *607 citing, 775 cited* | *607 citing, 775 cited* | |

Notes: cluster and heteroskedasticity robust standard errors in brackets; *** p < 0.01, ** p < 0.05, * p < 0.1

Consistent with the literature on other disciplines, being coauthors seems to be the most relevant determinant of citations within a pair of authors, with an estimated odds ratio of about 5.2 per coauthored paper in the pooled regressions, and 6.4 in the panel estimates. The interpretation of this coefficient is not straightforward, though, as being coauthors may imply that two economists work on similar or complementary topics, that they know each other's works better than they know third economists' works, and/or that they engage in strategic citing.

Various measures of proximity within a research group or scientific community appear as statistically significant correlates of dyadic citations. Working on similar topics, as measured by the cosine similarity of publications' metadata, has a conditional odds ratio of 2.9, while publishing in the same journal(s) of 3.3, and being affiliated with the same institution(s) of 1.9. Proximity in the policy debate is also found as a statistically significant predictor of the probability of citation, with each media outlet in which both *i* and *j* contributed in a same year increasing the odds of citation by roughly 5%.

When considering lagged values for all network connections (cols. 4-6), for the journals and media connections there are no substantial changes in the estimated coefficients and their statistical significance; for the former variable, lagged values emerge as significantly positive too. The estimated coefficient of coauthorship is lower in these new specifications, but coauthorship has a significantly positive impact in its lagged terms as well. Connections in terms of shared affiliations turns out to exert an impact not significantly different from zero (1 in terms of conditional odds ratios) in the specifications with lagged terms, whereas proximity in terms of topics only increases the odds of citation when taken with a time lag. Accordingly, the results seem to be not too different from the baseline estimations, and may suggest that citations arising from 'scientific' proximity (in terms of coauthorship and topics) may require more time to be made, possibly due to the time required to write or amend and then publish an article, whereas citations based on proximity in terms of academic journals and/or media outlets appear to depend on more immediate network relations.

On the whole, our estimates suggest that both scientific relations (e.g. similarity of topics) and social relations (e.g. working in the same institution or ideological proximity) determine economists' citation behavior. As discussed in section 2, while refraining here from defining scientific quality or merit, we interpret socially and politically driven citations as evidence of bias in citation behavior. However, the relevance of these factors would be smaller if one could expect that idiosyncratic biases cancel out in the aggregate. If that were case authors would be cited in proportion to their scientific impact within a certain field (where scientific impact would be measured exactly by what remains after biases cancel out). The only obstacle in equating citations with impact would then be that fields have different sizes, which could be dealt with by means of discipline or field normalization (Radicchi et al. 2008). Yet, this may not be the case if communities, institutions, and, notably, political positioning too vary in size (and thus at the very least their impact would need to be normalized).

To highlight this issue, following previous studies on other disciplines (Yan and Ding, 2009; Li et al., 2013; Uddin et al., 2013; Abbasi et al., 2014) we compute a number of network centrality

measures and estimate their impact on the total citations (net of self-citations) received in each year by Italy-based tenured academic economists.

### *5.1. Total citation counts*

We use three popular centrality measures adopted in the literature reviewed in section 3: degree centrality, betweenness centrality and closeness centrality. These measures were separately computed for each year for each network defined by the dyadic relations described above, namely, the co-authorship network, the network of institutional affiliations, the network of common journal authorships, the network of topics (cosine) similarity, and the network of contributors to national media outlets. Closeness centrality has been computed by attributing to non-existing paths (unconnected nodes) a length equal to the longest shortest path length observed in the network, plus one. Separately for each network, betweenness and closeness centrality have been normalized by dividing each value by the variable's yearly maximum, in order for them to take on values between 0 and 1 (degree centrality was not normalized because its values retain an intuitive interpretation).[11]

The intuition behind these measures is that degree centrality represents how many connections an author has; betweenness centrality measures how much an author could bridge different communities; and closeness centrality "how far" (expressed in number of links) the author is from all the other ones in the sample.

While extant literature mostly analyzes correlation matrices (reported in Table A2 in Appendix 2), we estimate a multivariate model of the total number of citations received in each year by an author, net of self-citations. As reported e.g. by Bornmann et al. (2008), in light of the count data nature of citations distributions, in the literature the Poisson or sometimes the Zero-Inflated Poisson distributions are the most popular functional forms for citation analysis.

Here we do not estimate zero-inflated models because there is no theoretical reason to assume that authors can be divided into two separate groups, i.e. there is no justification for assuming a process determining if an economist is cited at least once, different from the process determining how many times she is cited. Furthermore, in our sample the standard error of yearly citations to Italian tenured economists (25.8) is significantly higher than the mean (11.5), denoting over-dispersion. For these reasons, we estimate a conditional (fixed) effects panel Poisson model, with heteroskedasticity-robust standard errors obtained by the Huber-White sandwich estimator.

As a robustness check, we also estimate a pooled Poisson Pseudo-Maximum Likelihood (PPML) estimator with standard errors robust to clustering by author and to heteroskedasticity, as well as a negative binomial model with cluster- and heteroskedasticity-robust standard errors. As is well known, the latter is a parametric model aimed at tackling over-dispersion, such as is present in our data, by assuming an over-dispersion parameter with Gamma distribution with mean 1 and (estimated) variance $\alpha$. The former model requires more parsimonious hypotheses and it produces unbiased estimates of the conditional mean when

[11] Such a procedure is useful in order to reduce the correlation between centrality measures and, by changing the unit of measurement, to better highlight the sizes of the regression coefficients. It does not, however, qualitatively change the results of the estimates (further results are available from the author upon request).

tackling over-dispersion by means of heteroskedasticity-robust standard errors (Santos Silva and Tenreyro, 2006). However, it does not allow for estimating the distribution of the dependent variable, and as shown in Appendix 2, it leads to a considerably worse fit of the data. For these reasons the estimates from the conditional effects model are regarded here as superior, even though the model does not allow estimating time-invariant variables, and it attributes a considerable part of the overall variance to (unexplained) individual fixed effects.

As independent variables we include authors' degree centrality, betweenness centrality, and closeness centrality in the networks of co-authorships, affiliations, journal authorships, topics (with links based on cosine similarity) and entries on media outlets. Due to the high correlation of these measures (reported in Appendix 2), we consider the impact of each measure of centrality simultaneously across all networks but separately for the three different centrality measures (degree, betweenness, and closeness centrality). For the same reason, for closeness centrality we must exclude one network, i.e. that of common journals.

In all specifications control variables include authors' academic age, measured by the time since their first indexed publication, and authors' accumulated publications. The latter is included in logarithmic transformation because this way it can be considered as the authors' exposure to the possibility of being cited. When the associated coefficient is greater than zero (incidence rate ratio greater than one), the rate of yearly citations per published paper is not constant but rather it increases with the number of published papers, i.e. more prolific authors are more cited than the average.

For the conditional effects estimates, results are reported in Table 2. In this model, 16 authors had to be excluded from the estimates because in the period considered there is only one observation per author, and further 53 authors were excluded because the corresponding observations only exhibit zero outcomes (i.e. zero citations in all years). Therefore, these estimates are based on 3,445 observations for 706 authors.

The pooled PPML and the negative binomial models allow for the introduction of time-invariant regressors. Therefore in these models we further introduce authors' gender as a control variable. Furthermore, authors' prestige is often considered to be a relevant determinant of new citations, giving rise to the 'Matthew effect' described e.g. for the case of economics by Tol (2013). We proxy prestige by the authors' accumulated number of citations before the start of the period under analysis, that is, until 2010. However, due to a risk of possible endogeneity of such a variable, we also report all estimates excluding it. In all estimates, citations at the start of the period and academic age were normalized by subtracting the variables' mean value from each observation. Results for both the PPML and the negative binomial model are reported in Appendix 2.

In the baseline specification shown in table 2, columns 3-5, betweenness and closeness centrality in the network of Italian economists do not seem to exert a statistically significant impact on economists' total citation counts. In contrast, degree centrality both in the media and in the affiliations networks seems to positively impact on total citation counts. This result is partly in contrast with extant literature based on the notion of social capital. Usually assessing the impact on authors' productivity, betweenness centrality in the co-authorships network is often found to be an important factor (see e.g. Li et al., 2013). This result is usually interpreted in light of the fact that a network that channels information through the nodes with highest

betweenness centrality exhibits efficiency in lowering the number of links required to spread such information. Our estimates imply that the number of different people with whom an economist is connected, especially in those dimensions closer to the social sphere and further away from the scientific sphere (as could have been instead for connections in terms of topics or coauthorship) exerts a significant impact on authors' total citations. This is even more relevant when considering that the network of Italy-based economists is small with respect to the whole of Web of Science, from which total citation counts are taken.

Notably, for certain categories of economists the relevance of network connections, or lack thereof, appears to be higher than what aggregate values in the baseline estimations suggest. This is especially important for those categories of economists who, in a highly hierarchical discipline, stand at a disadvantage in terms of career opportunities with respect to the mainstream 'old male' tenured economist (Fourcade et al., 2015; Zacchia, 2017). When considering gender-specific impacts of network centrality (columns 6-8), betweenness centrality in the media network acquires statistical significance at the conventional 10% threshold. Similarly, when controlling for a heterogeneous impact of network centrality between older, established economists, and younger ones (respectively defined as those with an academic age above and below the median), the coefficients of closeness centrality both in the topics and in the affiliations networks increase in size and become statistically significant, as does betweenness centrality in the media and in the journals network (columns 9-11).

**Table 2. Determinants of authors' total citation counts: conditional effects Poisson model (incidence rate ratios)**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Baseline** | **Baseline** | **Degree cent.** | **Between-ness cent.** | **Closeness cent.** | **Degree cent.** | **Between-ness cent.** | **Closeness cent.** | **Degree cent.** | **Between-ness cent.** | **Closeness cent.** |
| Ln (publications) | 1.039** | 1.033*** | 1.030*** | 1.031*** | 1.032*** | 0.999 | 0.999 | 1.000 | 1.001 | 1.001 | 1.000 |
| | (0.017) | (0.011) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) |
| Academic age | | 1.169*** | 1.170*** | 1.169*** | 1.169*** | 1.169*** | 1.169*** | 1.169*** | 1.170*** | 1.169*** | 1.171*** |
| | | (0.012) | (0.011) | (0.012) | (0.012) | (0.011) | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) |
| Topics network | | | 1.000 | 1.024 | 0.995 | 1.000 | 1.054 | 1.009 | 1.000 | 1.019 | 1.098* |
| | | | (0.0001) | (0.041) | (0.046) | (0.0001) | (0.045) | (0.051) | (0.0001) | (0.053) | (0.062) |
| Affiliations network | | | 1.005* | 1.131 | 1.007 | 1.005** | 1.168 | 1.021 | 1.010*** | 1.351 | 0.864** |
| | | | (0.003) | (0.119) | (0.163) | (0.003) | (0.119) | (0.051) | (0.003) | (0.303) | (0.062) |
| Coauthorship network | | | 1.000 | 0.861 | 1.039 | 1.001 | 0.881 | 1.019 | 1.005 | 0.926 | 1.062 |
| | | | (0.014) | (0.119) | (0.072) | (0.015) | (0.119) | (0.037) | (0.020) | (0.180) | (0.057) |
| Journals network | | | 0.998 | 0.972 | | 0.998 | 0.947 | | 0.995* | 1.358* | |
| | | | (0.002) | (0.127) | | (0.002) | (0.125) | | (0.003) | (0.246) | |
| Media network | | | 1.002* | 1.327 | 0.993 | 1.002** | 1.468* | 0.986 | 1.003* | 1.685*** | 1.088 |
| | | | (0.001) | (0.336) | (0.048) | (0.001) | (0.340) | (0.052) | (0.002) | (0.337) | (0.097) |
| Int. woman * topics | | | | | | 1.000 | 0.943 | 0.967 | | | |
| | | | | | | (0.0002) | (0.084) | (0.095) | | | |
| Int. woman * affiliations | | | | | | 0.997 | 0.678 | 0.904 | | | |
| | | | | | | (0.006) | (0.237) | (0.107) | | | |
| Int. woman * coauthors | | | | | | 1.009 | 0.960 | 1.039 | | | |
| | | | | | | (0.037) | (0.568) | (0.089) | | | |
| Int. woman * journals | | | | | | 1.003 | 1.388 | | | | |
| | | | | | | (0.005) | (0.643) | | | | |
| Int. woman * media | | | | | | 0.997 | 0.280*** | 1.070 | | | |
| | | | | | | (0.002) | (0.100) | (0.113) | | | |
| Int. age * topics | | | | | | | | | 1.000 | 1.030 | 0.886 |
| | | | | | | | | | (0.0002) | (0.067) | (0.070) |
| Int. age * affiliations | | | | | | | | | 0.994 | 0.824 | 1.227** |
| | | | | | | | | | (0.004) | (0.203) | (0.112) |
| Int. age * coauthors | | | | | | | | | 0.992 | 0.983 | 0.953 |
| | | | | | | | | | (0.027) | (0.233) | (0.066) |
| Int. age * journals | | | | | | | | | 1.005 | 0.652* | |
| | | | | | | | | | (0.004) | (0.142) | |
| Int. age * media | | | | | | | | | 0.999 | 0.660 | 0.876 |
| | | | | | | | | | (0.002) | (0.246) | (0.087) |
| Observations | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 |
| Number of authors | 706 | 706 | 706 | 706 | 706 | 706 | 706 | 706 | 706 | 706 | 706 |

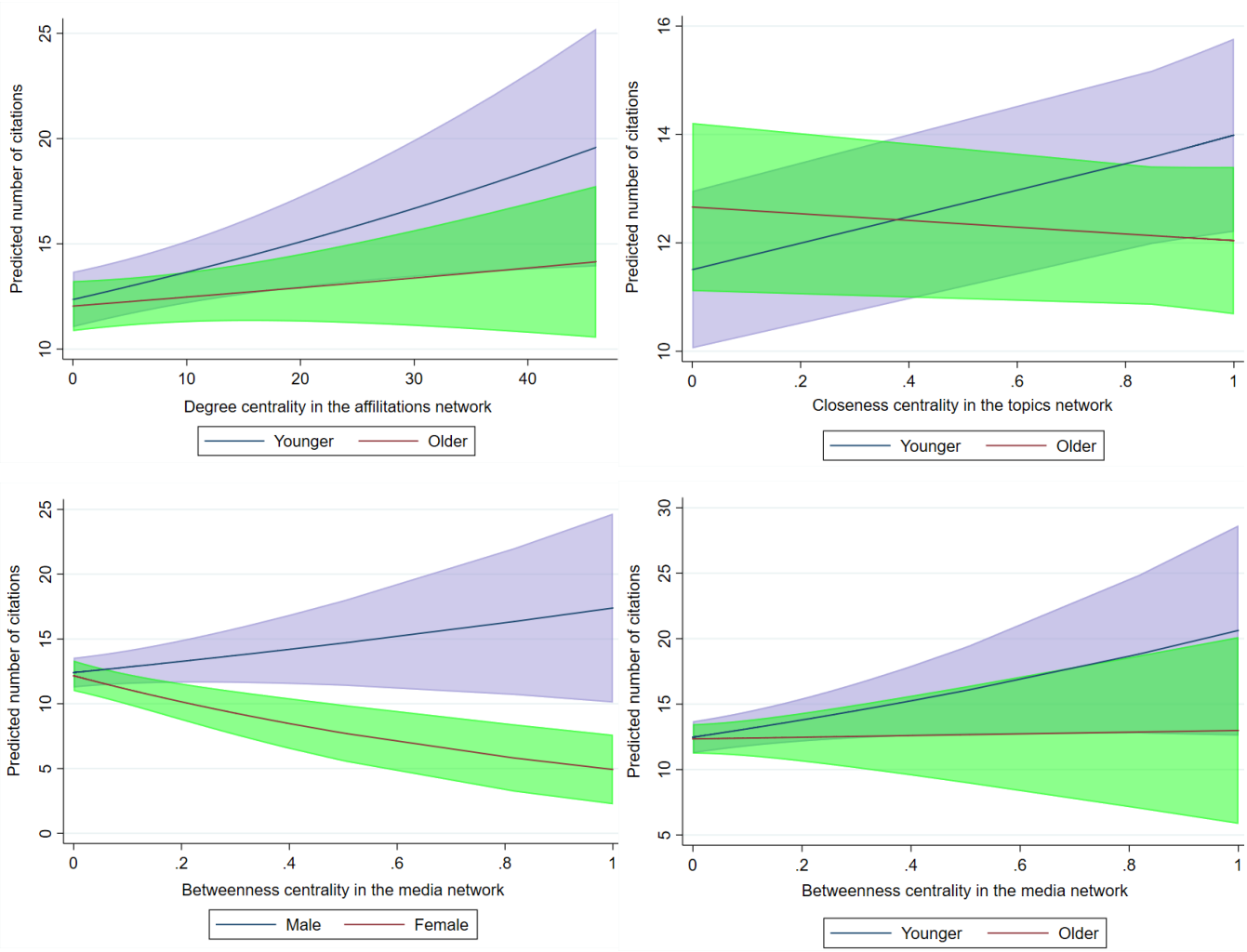
Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Accordingly, in the case of closeness centrality it appears that proximity to many other economists, in terms of topics/methods of investigation, is beneficial for younger authors, but it turns into a hurdle for citations to older economists, as shown in Fig. 1 (upper right panel). This latter finding possibly highlights a competitive disadvantage for 'eclectic' authors, and may reflect the well-known bibliometric disadvantage for multidisciplinary and interdisciplinary studies (see e.g. Rafols et al., 2012), though in our case the disadvantage emerges for authors who span different fields within the same discipline of economics.

Similarly, Fig. 1 highlights that betweenness centrality in the media network (lower panels) is associated with increasing citation counts for men, but not for women, and it is more beneficial for younger than for older authors. Similarly, having several colleagues in the same

institution(s) appears to be more beneficial for younger economists than for older ones (upper left panel), which may imply that as an economist progresses from the early stages of her career she becomes more widely known in the field at large.

**Figure 1. Predicted number of citations as a function of network centrality measures, by sex and academic age (average marginal effects of centrality with 95% confidence intervals)**




**Note:** marginal effects are computed from separate estimates for each graph, reported in columns 6-11 of table 2.

Concerning the robustness checks carried out by assuming different empirical models, we find that all specifications of the pooled PPML model (Table A4) correspond to a higher estimated impact of network centrality on authors' citation counts, especially when not controlling for authors' accumulated citations up to 2010. The impact of betweenness centrality becomes consistently positive and higher than what is estimated by the conditional effects model, while closeness centrality is confirmed as generally not significant, except for centrality in the media network.

The fact that centrality in the media network is found to exert a significantly positive impact in all specifications (including in the Negative Binomial model, Table A5) is especially interesting. Indeed, degree centrality in the media outlets network could in fact capture the effect of an author's skills or prestige, if media outlets selected the most 'talented' economists as regular columnists or contributors. Therefore, the interpretation of this coefficient may be difficult, as its positive sign could capture some unobserved characteristics that are not necessarily alien to scientific merit (and which in the conditional effects model are captured by the individual fixed effects). However, closeness centrality in the same network denotes the average 'ideological proximity' to the other economists in the sample, and betweenness centrality indicates her ability to bridge different communities in the media landscape. Therefore, these two measures are more clearly related to an author's positioning in the press/political landscape.

Ceteris paribus, we find that women receive fewer citations than men, as already found for several other disciplines (Abramo et al., 2015), but this impact is not statistically significant when controlling for authors' prestige (proxied by their accumulated citations at the beginning of the period) or when considering a negative binomial model, as shown in Table S7. Finally, in the negative binomial model estimates, all measures of centrality in the media network are generally confirmed to be positively correlated with authors' citation counts, as does centrality in the affiliations network (though in this case betweenness centrality, rather than degree centrality as in the conditional effect Poisson model). In this model, the estimated coefficient of betweenness centrality in the co-authorships network is considerably higher than in the other two models, and reaches the conventional statistical significance threshold of 5% when controlling for authors' prestige. In conclusion, robustness checks confirm the relevance of authors' centrality measures in several networks in shaping authors' citation counts, and even suggest a stronger impact of these variables than is implied by the conditional effects Poisson model.

In conclusion, citations among Italian economists appear to be driven by both scientific and social dynamics, including among the latter political-ideological proximity. Accordingly, the use of citation counts as measures of research quality establishes a distinct set of incentives for individual researchers in terms of fostering certain connections (and not others) both in the scientific and in the social field. For example, as an economist progresses in her career, pursuing a narrow disciplinary specialization may prove a promising strategy to maximize her expected citation counts, while a 'bipartisan' ideological attitude, bridging different political/media communities, seems especially convenient for men and younger authors. Survey evidence suggests that economists are, in part, already taking notice (Necker, 2014; Zacchia, 2017; May et al., 2018).

## 6. Conclusions

In this work we provide new evidence that citations reflect many more social trends than scientific impact alone, and that these additional elements do not become irrelevant when aggregating across publications to calculate bibliometric indexes at the individual level. Two main implications emerge.

On the one hand, citation indexes at the author level should be interpreted as reflecting a mix of both intellectual impact and social dynamics in non-obvious ways, and the use of citation

counts to rank authors does not lead to unambiguous measures of scientific impact. This finding, which was already highlighted for other disciplines, extends well beyond economics.
On the other hand, with specific respect to economics, in light of the relevance of social dynamics and especially political views in shaping citation behavior it may be necessary to question the efficiency of the discipline's organization in rewarding scientific achievement only. From this point of view, our analysis connects to the debate, reignited by the 2007-8 financial crisis and the ensuing "Great Recession", on the failures of contemporary economics (Bouchaud, 2008).

In trying to answer "the Queen's question", that is, why economists did not foresee the biggest crisis of the century, several authors consider institutional and social aspects as a relevant if not exclusive explanation. Considering American economics, Fourcade et al. (2015) document the closure of the economics discipline, which only cites other social sciences a negligible fraction of the times; its geographical concentration, with the rise of US-based journals vis-à-vis all the rest in the last few decades; and the institutional dominance of just five departments in the job market for new graduates and in the main national professional association.

Similarly, analyzing the curricula of all authors and editors of the top 4 general interest (US-based) economics journals, Colussi (2017) finds that the institutional dominance of six US departments shapes the selection of journal editors, who in turn, once selected, favor authors with whom they have network connections, such as current or former students and/or faculty colleagues.

By considering a larger dataset and focusing on the case of Italy, we show that the relevance of social ties in determining an economist's success in the 'citation game' is not limited to top-tier institutions in the leading country, but it is a widespread phenomenon with systematic consequences on authors' citation behavior. Further analyses will be required to assess how relevant social ties are in other disciplines.

## References


Abbasi A., Wigand R.T., Hossain L. (2014), "Measuring social capital through network analysis and its influence on individual performance", *Library & Information Science Research*, vol. 36, pp. 66-73.

Abramo, G., Cicero, T., C.A. D'Angelo (2015), "Should the research performance of scientists be distinguished by gender?", *Journal of Informetrics*, vol. 9, pp. 25-38.

Académie des Sciences, Leopoldina and Royal Society (2018), Statement by three national academies on good practice in the evaluation of researchers and research programmes, October 27, available at https://royalsociety.org/~/media/policy/Publications/2017/08-12-2017-royal-society-leopoldina-and-academie-des-sciences-call-for-more-support-for-research-evaluators.pdf

Adler R., Ewing J., Taylor P. (2008), *Citation Statistics. A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS)*, Joint IMU/ICIAM/IMS Committee on Quantitative Assessment of Research, available at http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf

Baccini A., De Nicolao G. (2016), "Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise", *Scientometrics*, vol. 108 n. 3, pp. 1651-1671.

Baldi S. (1998), "Normative versus Social Constructivist Processes in the Allocation of Citations: A Network- Analytic Model", *American Sociological Review*, vol. 63 n. 6, pp. 829-846.

Bertocchi G., Gambardella A., Jappelli T., Nappi C. A., Peracchi F. (2015), "Bibliometric evaluation vs. informed peer review: Evidence from Italy", *Research Policy*, vol. 44 n. 2, pp. 451-466.

Besancenot D., Huynh K., Serranito F. (2016), "Co-Authorship and Individual Research Productivity in Economics: Assessing The Assortative Matching Hypothesis", *Document de Recherche du Laboratoire d'Économie d'Orléans*, n. DRLEO 2015-16, Orléans.

Bornmann L., Mutz R., Neuhaus C., Daniel H.-D. (2008), "Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results", *Ethics in Science and Environmental Politics*, vol. 8, pp. 93-102.

Bouchaud J.-P. (2008), "Economics needs a scientific revolution", *Nature*, vol. 455 n. 30, p. 1181.

Brembs B., Button K., Munafò M. (2013), "Deep impact: unintended consequences of journal rank", *Frontiers in Human Neuroscience*, vol. 7, art. 291.

Cainelli G., Maggioni M.A., Uberti T.E., de Felice A. (2012), "Co-authorship and productivity among Italian economists", *Applied Economics Letters*, vol. 19 n. 6, pp. 1609-1613.

Camacho-Miñano M., Núñez-Nickel M. (2009), "The Multilayered Nature of Reference Selection", *Journal of the American Society for Information Science and Technology*, vol. 60 n. 4, pp. 754–777.

Cameron A.C., Gelbach J.G., and D.L. Miller (2008), "Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics*, vol. 90, pp. 414-427.

Cameron A.C., Gelbach J.G., and D.L. Miller (2011), "Robust Inference with Multi-Way Clustering", *The Journal of Business and Economic Statistics*, vol. 29 n. 2, pp. 238-249.

Case D.O., Miller J.B. (2011), "Do Bibliometricians Cite Differently From Other Scholars?", *Journal of the American Society for Information Science and Technology*, vol. 62 n. 3, pp. 421-432.

Colussi T. (2017), "Social Ties in Academia: A Friend is a Tresure", *The Review of Economics and Statistics*, forthcoming.

Corsi M., D'Ippoliti C., Lucidi F. (2011), "On the Evaluation of Economic Research: The case of Italy", *Economia Politica*, vol. 3, pp. 369-402.

D'Ippoliti C., Zacchia G. (2017), "On the Efficiency of Italian Universities: A Comment", *Italian Economic Journal*, vol. 3 n. 1, pp. 113-123.

Davis P.M. (2009), "Reward or persuasion? The battle to define the meaning of a citation", *Learned Publishing*, vol. 21, pp. 5-11.

Ductor L., Fafchamps M., Goyal S., van der Leij M.J. (2014), "Social Networks and Research Output", *The Review of Economics and Statistics*, vol. 96 n. 5, pp. 936-948.

Edwards M.A., Roy S. (2017), "Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition", *Environmental Engineering Science*, vol. 34 n. 1, pp. 51-61.

Fafchamps M., Goyal S., van der Leij M.J. (2010), "Matching and Network Effects", *Journal of the European Economic Association*, vol. 8 n. 1, pp. 203-231.

Fortunato S. *et al*. (2018), "Science of science", *Science*, n. 359, art. eaao0185, DOI: 10.1126/science.aao0185.

Fourcade M., Ollion E., Algan Y. (2015), "The Superiority of Economists", *Journal of Economic Perspectives*, vol. 29 n. 1, pp. 89-114.

Georg C.-P., Rose M.E. (2016a), "Mirror, Mirror on the Wall, Who is the Most Central of Them All?", *ERSA Working Paper*, n. 571, Economic Research Southern Africa, Cape Town.

Georg C.-P., Rose M.E. (2016b), "The Importance of Informal Intellectual Collaboration with Central Colleagues", Available at *SSRN*: https://ssrn.com/abstract=2877586.

Goyal S. (2016), "Networks in Economics. A Perspective on the Literature", in Bramoullé Y., Galeotti A., Rogers B. (eds.), *The Oxford Handbook of the Economics of Networks*, Oxford: Oxford University Press.

Goyal S., van der Leij M.J., Moraga-González J.-L. (2006), "Economics: An Emerging Small World", *Journal of Political Economy*, vol. 114 n. 2, pp. 403-412.

Graham B.S. (2017), "An Econometric Model of Network Formation with Degree Heterogeneity", *Econometrica*, forthcoming.

Helgadóttir O. (2016), "The Bocconi boys go to Brussels: Italian economic ideas, professional networks and European austerity", *Journal of European Public Policy*, vol. 23 n. 3, pp. 392-409.

Hicks D., Melkers J. (2012), "Bibliometrics as a Tool for Research Evaluation", in Link A., Vornatas N. (eds.), *Handbook on the Theory and Practice of Program Evaluation*, Cheltenham: Edward Elgar.

Hicks D., Potter J. (1991) "Sociology of scientific knowledge: a reflexive citation analysis or science dis- ciplines and disciplining science", *Social Studies of Science*, vol. 21, pp. 459-501.

Jackson M.O., Rogers B.W., Zenou Y. (2017), "The Economic Consequences of Social-Network Structure", *Journal of Economic Literature*, vol. 55 n. 1, pp. 49-95.

Johnson and Oppenheim, 2007;

King D.A. (2004), "The scientific impact of nations", *Nature*, vol. 430, pp. 311-316.

King M.M., Bergstrom C.T., Correll S.J., Jacquet J., West J.D. (2016), "Men set their own

cites high: Gender and self-citation across fields and over time", *arXiv*, n. 1607.00376, available at https://arxiv.org/abs/1607.00376

Krampen G., Becker R., Wahner U., Montada L. (2007), "On the validity of citation counting in science evaluation: Content analyses of references and citationsin psychological publications", *Scientometrics*, vol. 71 n. 2, pp. 191-202.

Letchford A., Moat H.S., Preis T. (2015) "The advantage of short paper titles", *Royal Society Open Science*, vol. 2, art. 150266.

Li E.Y., Liao C.H., Yen H.R. (2013), "Co-authorship networks and research impact: A social capital perspective", *Research Policy*, vol. 42, pp. 1515-1530.

Macdonald S., Kam J. (2010), "Counting footnotes: Citability in management studies", *Scandinavian Journal of Management*, vol. 26, pp. 189-203.

MacRoberts M.H., MacRoberts B.R. (1996), "Problems of citation analysis", *Scientometrics*, vol. 36 n. 3, pp. 435-444.

Mählck P., Persson O. (2000), "Socio-bibliometric mapping of intra-departmental networks", *Scientometrics*, vol. 49, pp. 81-91.

Mancini P. (2013), "The Italian public sphere: a case of dramatized polarization", *Journal of Modern Italian Studies*, vol. 18 n. 3, pp. 335-347.

Martin B.R., Irvine J. (1983), "Assessing Basic Research", *Research Policy*, vol. 12, pp. 61-90.

May A.M., McGarvey M.G., Kucera D. (2018), "Gender and European Economic Policy: A Survey of the Views of European Economists on Contemporary Economic Policy", *Kyklos*, vol. 71, pp. 162-183.

Moed E.F. (2005), *Citation Analysis in Research Evaluation*, Dordrecht: Springer.

Molina J.A., Alcolea A., Ferrer A., Iñiguez D., Rivero A., Ruiz G., Tarancón A. (2016), "Co-authorship and Academic Productivity in Economics: Interaction Maps from the Complex Networks Approach", *IZA Discussion Paper*, n. 10008, Forschungsinstitutzur Zukunft der Arbeit, Bonn.

Necker S. (2014), "Scientific misbehavior in economics", *Research Policy*, vol. 43 n. 10, pp. 1747-1759.

Oswald A.J. (2007). "An Examination of the Reliability of Prestigious Scholarly Journals: Evidence and Implications for Decision-Makers", *Economica*, vol. 74 n. 293, pp. 21-31.

Perc M. (2014), "The Matthew effect in empirical data", *Journal of the Royal Society Interface*, vol. 11 n. 20140378.

Pfeiffer T., Hoffmann R. (2007) "Temporal patterns of genes in scientific publications", *Proceedings of the National Academy of Sciences USA*, vol. 104, pp. 12052-12056.

Radicchi F., Fortunato S., Castellano C. (2008), "Universality of citation distributions: towards an objective measure of scientific impact", *Proceedings of the National Academy of Sciences USA*, vol. 105 n. 45, pp. 17268–17272.

Rafols I., Leydesdorff L., O'Hare A., Nightingale P., Stirling A. (2012), "How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management", *Research Policy*, vol. 41, pp. 1262-1282.

Santos Silva J.M.C., Tenreyro S. (2006), "The Log of Gravity", *The Review of Economics and Statistics*, vol. 88 n. 4, pp. 641-658.

Sud P., Thelwall M. (2014), "Evaluating altmetrics", *Scientometrics*, vol. 98, pp. 1131-1143.

Tol R.S.J. (2013), "The Matthew effect for cohorts of economists", *Journal of Informetrics*, vol. 7, pp. 522–527.

Uddin S., Hossain L., Rasmussen K. (2013), "Network Effects on Scientific Collaborations", *PLOS ONE*, vol. 8 n. 2, art. e57546.

Wallace M.L., Larivière V., Gingras Y. (2012), "A Small World of Citations? The Influence of Collaboration Networks on Citation Practices", *PLOS ONE*, vol. 7 n. 3, art. e33339.

Wang D., Song C., Barabási A.-L. (2013), "Quantifying long-term scientific impact", *Science*, vol. 342, pp. 127-132.

White H.D. (2001), "Authors as citers over time", *Journal of the American Society for Information Science and Technology*, vol. 52 n. 2, pp. 87-108.

White H.D., Wellman B., Nazer N. (2004), "Does Citation Reflect Social Structure? Longitudinal Evidence From the 'Globenet' Interdisciplinary Research Group", *Journal of the American Society for Information Science and Technology*, vol. 55 n. 2, pp. 111-126.

White M.D., Wang P. (1997), "A Qualitative Study of Citing Behavior: Contributions, Criteria, and Metalevel Documentation Concerns", *The Library Quarterly*, vol. 67 n. 2, pp. 122-154.

Wilhite A.W., Fong E.A. (2012), "Coercive Citation in Academic Publishing", *Science*, n. 335, pp. 542-543.

Yan E., Ding Y. (2009), "Applying Centrality Measures to Impact Analysis: A Coauthorship Network Analysis", *Journal of the American Society for Information Science and Technology*, vol. 60 n. 10, pp. 2107-2118.

Zacchia G. (2016), "Paolo Sylos Labini: Reflections on a Classical Economist", *PSL Quarterly Review*, vol. 69, pp. 199-208.

Zacchia G. (2017), "Diversity in Economics: A Gender Analysis of Italian Academic Production", *INET Working Paper*, n. 2017m, available at https://www.ineteconomics.org/uploads/papers/WP_61-Zacchia-Diversity-Final.pdf

## Appendix 1 – Sources and methods

### *A1.1. Mass media outlets considered*

The following 68 print newspapers and magazines were considered for the analysis:

*Il Corriere della Sera, la Repubblica, la Stampa, Il Sole 24 Ore, Il Messagero, Il Giornale, Il Resto del Carlino, Libero Quotidiano, L'Unità, Avvenire, Il Fatto Quotidiano, Il Foglio, Il Manifesto, Tempo, Il Riformista, Il Mattino, Panorama, l'Espresso, Il Secolo d'Italia, Il Giorno, Italia Oggi, Left, L'Osservatore Romano, Liberazione, Europa, La Gazzetta del Mezzogiorno, Il Secolo XIX, L'Opinione, Il Mondo, Finanza e Mercati, Internazionale, Diario, Pubblico, Famiglia Cristiana, L'Avanti, Milano Finanza (MF), Il Gazzettino, Anna (Annabella), Il Corriere Adriatico, Il Corriere del Veneto, Il Corriere della Comunicazione, Il Giornale di Sicilia, Gli Altri, Il Garantista, Oggi, Giustizia, La Discussione, Terra, Popolo, Roma, Russia Beyond the Headlines, La Padania, Vanity Fair, La Rinascita della Sinistra, L'Altro, Via Sarfatti 25, Il Corriere del Mezzogiorno, Il Campanile, L'Indipendente, Il Giornale della Toscana, La Nazione, Civiltà, La Voce Repubblicana, Il Mulino, Formiche, Tempi, Economy, La Notizia;*

as well as the following ten online magazines and blogs:

*NoiseFromAmerika.org, laVoce.info, nelMerito.com, inGenere.it, Sbilanciamoci.info, EconomiaePolitica.it, neoDemos.info, voxEU.org, SviluppoFelice.wordpress.com, Keynesblog.com.*

### *A1.2. List of synonyms for the sake of content analysis*

Based on the 400 most common terms in the articles' metadata, the following terms were considered – strictly for the sake of the computation of the cosine similarity only – as synonyms:

| **Lemmas** | **Aggregated with** |
|---|---|
| accounting | accountability |
| africans | africa |
| americans | america |
| american | america |
| bank-and-time | bank |
| bank-backed | bank |
| bank-based | bank |
| bank-client | bank |
| bank-enterprise | bank |
| banker | bank |
| bank-firm | bank |
| bank-fund | bank |
| banking | bank |
| bank-lending | bank |
| bank-manager | bank |
| bankruptcy | bank |
| interbank | bank |
| bank-specific | bank |
| bank-to-bank | bank |
| banque | bank |

| | |
|---|---|
| employability | labour |
| employee | labour |
| employee-level | labour |
| employer | labour |
| employer-employee | labour |
| employer-provided | labour |
| employers-invest | labour |
| employment | labour |
| employment-based | labour |
| employment-enhancing | labour |
| employment-oriented | labour |
| employment-productivity | labour |
| employment-related | labour |
| employment-share | labour |
| labor | labour |
| labor-cost | labour |
| labor-intensive | labour |
| labor-market | labour |
| labour-father | labour |
| labour-intensive | labour |

| | |
|---|---|
| labour-market | labour |
| labour-saving | labour |
| unemployment | labour |
| unemployment-vacancy | labour |
| unemployment-vacancies | labour |
| wage | labour |
| wage-cost | labour |
| wage-based | labour |
| wage-flexibility | labour |
| wage-moderation | labour |
| wage-neutral | labour |
| wage-oriented | labour |
| wage-profits-pensions | labour |
| wages-capital-intensive | labour |
| wage-setting | labour |
| wages-labor-intensive | labour |
| worker | labour |
| worker-firm | labour |

| | |
|---|---|
| worker-job | labour |
| workers-consumers | labour |
| work-experienced | labour |
| workforce | labour |
| work-force | labour |
| work-hours | labour |
| working | labour |
| working-age | labour |
| workload | labour |
| workplace | labour |
| work-profile | labour |
| work-related | labour |
| work-sharing | labour |
| eu | europe |
| eu-12 | euro |
| eu15 | euro |
| eu-15 | euro |
| eu25 | europe |
| eu27 | europe |
| eu-27 | europe |
| eurobarometer | europe |
| euroland | euro |
| euromed | europe |
| euro-mediterranean | europe |
| europaregion | europe |
| european | europe |
| europeanism | europe |
| europeanization | europe |
| european-wide | europe |
| eurosystem | euro |
| eurozone | euro |
| euro-dollar | euro |
| euro-usd | euro |
| finance | finance |
| finance-growth | finance |
| finances | finance |
| financial | finance |
| financialisation | finance |
| financialization | finance |
| financially | finance |
| financially-developed | finance |
| financialmarket | finance |
| financial-market | finance |
| financial-network | finance |
| financial-related | finance |
| household-level | household |
| household-specific | household |
| family | household |
| family-level | household |
| italia | italy |
| italian | italy |
| italian | italy |
| italiana | italy |
| italians | italy |
| italien | italy |
| policies-practices | policy |
| policy | policy |
| Policy | policy |
| policy.classification | policy |
| policy-decision | policy |
| policy-independent | policy |

| | |
|---|---|
| policy-induced | policy |
| policymaker | policy |
| policy-maker | policy |
| policymaking | policy |
| policy-making | policy |
| policy-motivated | policy |
| policy-oriented | policy |
| policy-relevant | policy |
| academic | academia |
| accessibility | access |
| african | africa |
| agri-food | agriculture |
| agricultural | agriculture |
| airways | airline |
| alitalia | airline |
| allocate | allocation |
| antidump | antidumping |
| artist | art |
| asymmetrical | asymmetry |
| behavioral | behavior |
| borrowing | borrower |
| capitalist | capitalism |
| chinese | china |
| consumption-production | consumption |
| consumption-expenditure | consumption |
| consumer | consumption |
| consumption-real | consumption |
| contingency | contingent |
| converge | convergence |
| cost-plus | cost |
| country-specific | country |
| creditor | credit |
| criminal | crime |
| currency-risk | currency |
| debtor | debt |
| debt-targeting | debt |
| debt-gdp | debt |
| decentralize | decentralisation |
| deficit-gdp | deficit |
| delinquent | delinquency |
| democratic | democracy |
| duopolistic | duopoly |
| dynamization | dynamic |
| dynamics | dynamic |
| earning | earnings |
| entrepreneurial | entrepreneur |
| environmental | environment |
| environmental-economic | environment |
| environmentally | environment |
| epistemological | epistemology |
| evolutionary | evolution |
| experimenter | experiment |
| exploitative | exploitation |
| exporter | export |
| export-share-weighted | export |
| fdi-related | fdi |
| quantitatively | quantitative |

| | |
|---|---|
| rationalize | rational |
| reallocate | reallocation |
| regional-national | region |
| regional | region |
| region-specific | region |
| religious | religiosity |
| remunerate | remuneration |
| rewarding | reward |
| risk-sharing | risk |
| risk-adjusted | risk |
| risk-management | risk |
| risk-seeking | risk |
| rule-based | rule |
| salaried | salary |
| scoring | score |
| sector-oriented | sector |
| sectorial | sector |
| sectoral | sector |
| self-interested | self-interest |
| shareholding | shareholder |
| simulation | simulate |
| sizeable | size |
| skill-productivity | skill |
| skilled-labour | skill |
| skill-replacing | skill |
| sraffa | sraffian |
| stabilization | stability |
| sub-national | subnational |
| technology-augmented | technology |
| technology-based | technology |
| technological | technology |
| time-span | time |
| time-variation | time |
| tourism-led | tourism |
| tourist | tourism |
| trade-trade | trade |
| trade-related | trade |
| trade-based | trade |
| trader | finance |
| trading | finance |
| transportation | transport |
| u.k. | uk |
| britain | uk |
| u.s. | usa |
| value-at-risk | finance |
| volatile | volatility |
| voting | vote |
| voter | vote |
| waste-policy | waste |
| welfare-increasing | welfare |
| welfare-reducing | welfare |
| welfare-dominates | welfare |
| youth | young |
| climate | environment |
| emission | environment |
| family | household |
| child | household |
| households | household |
| preference | utility |
| estate | housing |
| offshoring | offshore |

| | |
|---|---|
| offshor | offshore |
| belief | expectation |
| lazio | italy |
| rome | italy |
| kenyans | kenya |
| new-keynesian | keynesian |
| post-keynesian | keynes |
| knowledge-intensive | knowledge |
| labour-market | labor |
| legislator | politics |
| legislature | politics |
| lobbying | lobby |
| macroeconomics | macroeconomic |
| managers | manager |
| manufacturing | manufacture |
| market-specific | market |
| married | marriage |
| marx | marxian |
| methodological | method |
| methodologically | method |
| migrant | migration |
| model-based | model |
| modelling | model |
| monetary-fiscal | monetary |
| monopolist | monopoly |
| neighborhood | neighbor |
| neo-classical | neoclassical |
| neoliberalism | neoliberal |
| networking | network |
| netwok-based | network |
| now-casting | nowcast |
| nowcasting | nowcast |
| older | old |
| optimality | optimal |
| outsider | outside |
| outsourcing | outsource |
| owner | ownership |
| paradigmatic | paradigm |

| | |
|---|---|
| parametrization | parameter |
| pareto-dominated | pareto |
| paretian | pareto |
| peculiarity | peculiar |
| permanent-temporary | permanent |
| politics-business | politics |
| politician | politics |
| pollute | pollution |
| predictive | prediction |
| probabilistic | probability |
| processing | process |
| pro-cyclical | procyclical |
| producer | production |
| profit-seeking | profit |
| providers | provider |
| prudencial | prudent |
| quality-ladder | quality |
| quality-oriented | quality |
| financially | finance |
| financial-market | finance |
| finance-growth | finance |
| financial | finance |
| finances | finance |
| firms-countries | firm |
| forecasting | forecast |
| forecaster | forecast |
| forecasters | forecast |
| formalise | formalize |
| freedom-corruption | freedom |
| friendship | friend |
| generalize | generalisation |
| generous | generosity |
| g-20 | global |
| g3 | global |
| globally | global |
| g7 | global |
| globalisation | global |
| governing | govern |

| | |
|---|---|
| governance | govern |
| greenhouse | environment |
| growth-oriented | growth |
| hedging | hedge |
| hi-tech | ict |
| high-tech | ict |
| human-generated | human |
| human-induced | human |
| imperfection | imperfect |
| individual-level | individual |
| individually-owned | individual |
| individually | individual |
| industrialized | industrialize |
| industry-region | industry |
| industry-wide | industry |
| industrial | industry |
| inflation-targeting | inflation |
| inflation-averse | inflation |
| inflation-unemployment | inflation |
| inflation-targeting | inflation |
| inflation-averse | inflation |
| inflation-unemployment | inflation |
| infrastructural | infrastructure |
| innovation-driven | innovation |
| innovator | innovation |
| innovative | innovation |
| innovating | innovation |
| interdependency | interdependence |
| inter-generational | intergenerational |
| investor | investing |

# Appendix 2 – Additional results

## Table A1 – Dyadic relations: descriptive statistics for each pair in the sample

| | Obs. | N. of connections | | Min. | Max. |
|---|---|---|---|---|---|
| At least one citation in *t* | 1,148,348 | 0.025% | | 0 | 1 |
| Common journals in *t* | 1,148,348 | 2.029% | | 0 | 7 |
| Coauthorships in *t* | 1,148,348 | 0.124% | | 0 | 6 |
| Common affiliations in *t* | 1,148,348 | 0.946% | | 0 | 2 |
| Same media outlets in *t* | 1,148,348 | 3.256% | | 0 | 56 |
| | **Obs.** | **Mean** | **Std. Dev.** | **Min.** | **Max.** |
| Citations in *t* | 1,148,348 | 0.0003 | 0.0223 | 0 | 6 |
| Cosine similarity of metadata up to *t* | 1,148,348 | 3.27% | 6.36% | 0 | 1 |
| Number of connections in *t* | 1,148,348 | 0.39 | 0.53 | 0 | 5 |
| Networks in which there is a connection in *t* | 1,148,348 | 0.10 | 0.73 | 0 | 56 |
| Publications by cited author, up to *t* | 1,148,348 | 3.05 | 3.78 | 1 | 29 |
| Publications by citing author, in t | 1,148,348 | 1.58 | 0.98 | 1 | 9 |

## Table A2 – Descriptive statistics at the author level

| | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Total citations (exc. self-citations) in *t* | 3700 | 11.53 | 25.79 | 0 | 296 |
| Publications up to *t* | 3700 | 8.83 | 11.24 | 1 | 105 |
| Academic age | 3700 | 10.71 | 7.79 | 0 | 30 |
| Citations up to 2010 | 3700 | 32.61 | 127.86 | 0 | 1590 |
| Female | 3700 | 26.4% | | | |
| Degree centrality: topics network | 3700 | 114.65 | 141.63 | 0 | 331 |
| Degree centrality: affiliations network | 3700 | 2.89 | 5.67 | 0 | 46 |
| Degree centrality: coauthorship network | 3700 | 0.34 | 0.73 | 0 | 8 |
| Degree centrality: journals network | 3700 | 2.99 | 5.75 | 0 | 44 |
| Degree centrality: media network | 3700 | 2.21 | 7.82 | 0 | 63 |
| Betweenness centrality: topics network | 3700 | 0.25 | 0.32 | 0 | 1 |
| Betweenness centrality: affiliations network | 3700 | 0.01 | 0.06 | 0 | 1 |
| Betweenness centrality: coauthorship net. | 3700 | 0.01 | 0.04 | 0 | 1 |
| Betweenness centrality: journals network | 3700 | 0.01 | 0.06 | 0 | 1 |
| Betweenness centrality: media network | 3700 | 0.004 | 0.03 | 0 | 1 |
| Closeness centrality: topics network | 3700 | 0.39 | 0.48 | 0 | 1 |
| Closeness centrality: affiliations network | 3700 | 0.25 | 0.36 | 0 | 1 |
| Closeness centrality: coauthorship network | 3700 | 0.18 | 0.33 | 0 | 1 |
| Closeness centrality: journals network | 3700 | 0.28 | 0.36 | 0 | 1 |
| Closeness centrality: media network | 3700 | 0.1 | 0.26 | 0 | 1 |

## Table A3 – Correlation between various network centrality measures

| | Degree centrality | | | | | Betweenness centrality | | | | | Closeness centrality | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Topics** | **Affil.** | **Coaut.** | **Jour.** | **Media** | **Topics** | **Affil.** | **Coaut.** | **Jour.** | **Media** | **Topics** | **Affil.** | **Coaut.** | **Jour.** |
| **Degree centrality** | | | | | | | | | | | | | | |
| Affiliations | 0.5995 | | | | | | | | | | | | | |
| | [0.000] | | | | | | | | | | | | | |
| Couthorship | 0.5424 | 0.4122 | | | | | | | | | | | | |
| | [0.000] | [0.000] | | | | | | | | | | | | |
| Journals | 0.5903 | 0.4424 | 0.5633 | | | | | | | | | | | |
| | [0.000] | [0.000] | [0.000] | | | | | | | | | | | |
| Media | 0.1350 | 0.1731 | 0.1515 | 0.1202 | | | | | | | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | | | | | | | | | | |
| **Betweenness centrality** | | | | | | | | | | | | | | |
| Topics | 0.9604 | 0.5718 | 0.5256 | 0.5678 | 0.1319 | | | | | | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | | | | | | | | |
| Affiliations | 0.2394 | 0.4899 | 0.1699 | 0.2243 | 0.0777 | 0.2241 | | | | | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | | | | | | | |
| Couthorship | 0.1371 | 0.1299 | 0.5106 | 0.3153 | 0.1015 | 0.1345 | 0.0899 | | | | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | | | | | | |
| Journals | 0.2795 | 0.3004 | 0.3948 | 0.6534 | 0.1284 | 0.2660 | 0.2344 | 0.3924 | | | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | | | | | |
| Media | 0.0902 | 0.1042 | 0.0948 | 0.0886 | 0.5268 | 0.0821 | 0.0148 | 0.0248 | 0.0664 | | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.367] | [0.131] | [0.000] | | | | | |
| **Closeness centrality** | | | | | | | | | | | | | | |
| Topics | 0.9923 | 0.5949 | 0.5447 | 0.5973 | 0.1378 | 0.9505 | 0.2288 | 0.1339 | 0.2748 | 0.0973 | | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | | | |
| Affiliations | 0.7669 | 0.7305 | 0.5437 | 0.5628 | 0.1337 | 0.7456 | 0.2662 | 0.1605 | 0.3023 | 0.0836 | 0.7847 | | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | | |
| Couthorship | 0.6196 | 0.4759 | 0.8373 | 0.5613 | 0.1225 | 0.5980 | 0.1841 | 0.1938 | 0.3075 | 0.0755 | 0.6241 | 0.6104 | | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | | |
| Journals | 0.8721 | 0.5775 | 0.5820 | 0.6936 | 0.1360 | 0.8434 | 0.2239 | 0.1457 | 0.3127 | 0.0968 | 0.8894 | 0.7710 | 0.6822 | |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | |
| Media | 0.0312 | 0.0690 | 0.0769 | 0.0429 | 0.7567 | 0.0358 | 0.0379 | 0.0652 | 0.0597 | 0.3354 | 0.0311 | 0.0456 | 0.0487 | 0.0386 |
| | [0.057] | [0.000] | [0.000] | [0.009] | [0.000] | [0.029] | [0.021] | [0.000] | [0.000] | [0.000] | [0.050] | [0.006] | [0.003] | [0.019] |

*Note*: p-values in brackets.

**Table A4 – Determinants of authors' yearly citations: Poisson Pseudo-Maximum Likelihood model (incidence rate ratios)**

| | (1) | (2) | (3) | (4) | (6) | (5) | (7) Degree centrality | (8) Degree centrality | (9) Between-ness cent. | (10) Between-ness cent. | (11) Closeness centrality | (12) Closeness centrality |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ln (acc. publications) | 1.461*** | 1.466*** | 1.463*** | 1.239*** | 1.451*** | 1.235*** | 1.374*** | 1.209*** | 1.369*** | 1.196*** | 1.391*** | 1.207*** |
| | (0.094) | (0.087) | (0.090) | (0.079) | (0.087) | (0.078) | (0.077) | (0.067) | (0.073) | (0.067) | (0.075) | (0.064) |
| Academic age | | 1.070*** | 1.068*** | 1.036*** | 1.066*** | 1.034*** | 1.064*** | 1.034*** | 1.066*** | 1.034*** | 1.064*** | 1.033*** |
| | | (0.007) | (0.007) | (0.006) | (0.007) | (0.006) | (0.007) | (0.006) | (0.006) | (0.006) | (0.007) | (0.006) |
| Woman | | | | | 0.724* | 0.833 | 0.738* | 0.850 | 0.736* | 0.848 | 0.748 | 0.870 |
| | | | | | (0.136) | (0.152) | (0.133) | (0.152) | (0.136) | (0.153) | (0.137) | (0.156) |
| Topics network | | | | | | | 1.000 | 1.000 | 1.031 | 1.193 | 0.913 | 1.174 |
| | | | | | | | (0.0004) | (0.0004) | (0.111) | (0.137) | (0.130) | (0.132) |
| Affiliations network | | | | | | | 1.015** | 1.014* | 2.905*** | 3.571*** | 1.254 | 0.943 |
| | | | | | | | (0.007) | (0.008) | (0.881) | (1.139) | (0.215) | (0.119) |
| Couthorship network | | | | | | | 1.074* | 0.982 | 1.541* | 1.187 | 1.166 | 1.050 |
| | | | | | | | (0.044) | (0.046) | (0.378) | (0.772) | (0.169) | (0.121) |
| Journals network | | | | | | | 0.995 | 0.996 | 2.001*** | 0.457** | | |
| | | | | | | | (0.005) | (0.006) | (0.532) | (0.152) | | |
| Media network | | | | | | | 1.021*** | 1.013** | 9.183*** | 4.726** | 1.890*** | 1.759*** |
| | | | | | | | (0.006) | (0.005) | (6.728) | (3.704) | (0.442) | (0.338) |
| Linear time trend | | | 1.079*** | 1.114*** | 1.082*** | 1.116*** | 1.083*** | 1.122*** | 1.073*** | 1.122*** | 1.088*** | 1.114*** |
| | | | (0.015) | (0.013) | (0.015) | (0.013) | (0.015) | (0.015) | (0.015) | (0.015) | (0.015) | (0.015) |
| Citations up to 2010 | | | | 1.002*** | | 1.002*** | | 1.002*** | | 1.002*** | | 1.002*** |
| | | | | (0.0003) | | (0.0003) | | (0.0003) | | (0.0003) | | (0.0003) |
| Constant | 5.702*** | 4.861*** | 3.844*** | 4.460*** | 4.166*** | 4.666*** | 4.153*** | 4.366*** | 4.457*** | 4.489*** | 3.808*** | 4.212*** |
| | (0.929) | (0.789) | (0.671) | (0.707) | (0.715) | (0.707) | (0.647) | (0.611) | (0.719) | (0.624) | (0.626) | (0.622) |
| Observations | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 |

Note: cluster robust standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

**Table A5 – Determinants of authors' yearly citations: Negative Binomial model (incidence rate ratios)**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Degree cent. | Degree cent. | Between-ness cent. | Between-ness cent. | Closeness cent. | Closeness cent. |
| Ln (publications) | 1.370*** | 1.339*** | 1.334*** | 1.318*** | 1.310*** | 1.320*** | 1.270*** | 1.261*** | 1.262*** | 1.255*** | 1.276*** | 1.273*** |
| | (0.069) | (0.061) | (0.062) | (0.051) | (0.050) | (0.060) | (0.047) | (0.037) | (0.047) | (0.038) | (0.046) | (0.039) |
| Academic age | | 1.100*** | 1.098*** | 1.033*** | 1.032*** | 1.096*** | 1.091*** | 1.032*** | 1.094*** | 1.033*** | 1.092*** | 1.031*** |
| | | (0.010) | (0.010) | (0.008) | (0.008) | (0.010) | (0.010) | (0.008) | (0.010) | (0.008) | (0.010) | (0.008) |
| Woman | | | | | 0.827 | 0.803 | 0.823 | 0.831 | 0.813 | 0.834 | 0.837 | 0.836 |
| | | | | | (0.103) | (0.154) | (0.160) | (0.104) | (0.155) | (0.104) | (0.162) | (0.105) |
| Topics network | | | | | | | 1.000 | 1.000 | 0.966 | 1.061 | 0.951 | 0.918 |
| | | | | | | | (0.000) | (0.000) | (0.103) | (0.0935) | (0.117) | (0.0933) |
| Affiliations network | | | | | | | 1.011 | 1.010 | 3.562*** | 2.312*** | 1.101 | 1.163 |
| | | | | | | | (0.007) | (0.006) | (1.424) | (0.719) | (0.157) | (0.132) |
| Couthorship network | | | | | | | 1.048 | 1.062 | 1.978 | 2.243** | 1.141 | 1.132 |
| | | | | | | | (0.043) | (0.041) | (0.834) | (0.897) | (0.123) | (0.095) |
| Journals network | | | | | | | 0.997 | 0.999 | 1.715 | 1.567 | | |
| | | | | | | | (0.006) | (0.005) | (0.667) | (0.640) | | |
| Media network | | | | | | | 1.023*** | 1.014*** | 36.66** | 5.086 | 1.872*** | 1.409** |
| | | | | | | | (0.006) | (0.004) | (57.85) | (6.857) | (0.402) | (0.194) |
| Citations up to 2010 | | | | 1.006*** | 1.006*** | | | 1.005*** | | 1.005*** | | 1.005*** |
| | | | | (0.001) | (0.001) | | | (0.001) | | (0.001) | | (0.001) |
| Linear time trend | | | 1.096*** | 1.179*** | 1.179*** | 1.096*** | 1.100*** | 1.176*** | 1.094*** | 1.172*** | 1.101*** | 1.182*** |
| | | | (0.017) | (0.015) | (0.015) | (0.017) | (0.018) | (0.016) | (0.016) | (0.015) | (0.018) | (0.015) |
| Ln (alpha) | 2.471*** | 2.071*** | 2.054*** | 1.623*** | 1.616*** | 2.045*** | 1.995*** | 1.593*** | 2.021*** | 1.605*** | 2.010*** | 1.602*** |
| | (0.141) | (0.121) | (0.123) | (0.090) | (0.090) | (0.127) | (0.127) | (0.089) | (0.126) | (0.090) | (0.126) | (0.089) |
| Constant | 6.425*** | 5.453*** | 4.112*** | 2.791*** | 2.956*** | 4.410*** | 4.240*** | 2.968*** | 4.552*** | 3.076*** | 4.108*** | 2.868*** |
| | (0.905) | (0.683) | (0.633) | (0.279) | (0.304) | (0.652) | (0.615) | (0.296) | (0.661) | (0.309) | (0.622) | (0.296) |
| Observations | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 | 3,700 |

Note: cluster robust standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

**Figure A1.** The connected components in the networks of coauthorship (A), common institutional affiliations (B), common media outlets (C), and in the union of all three (D), for year 2011.

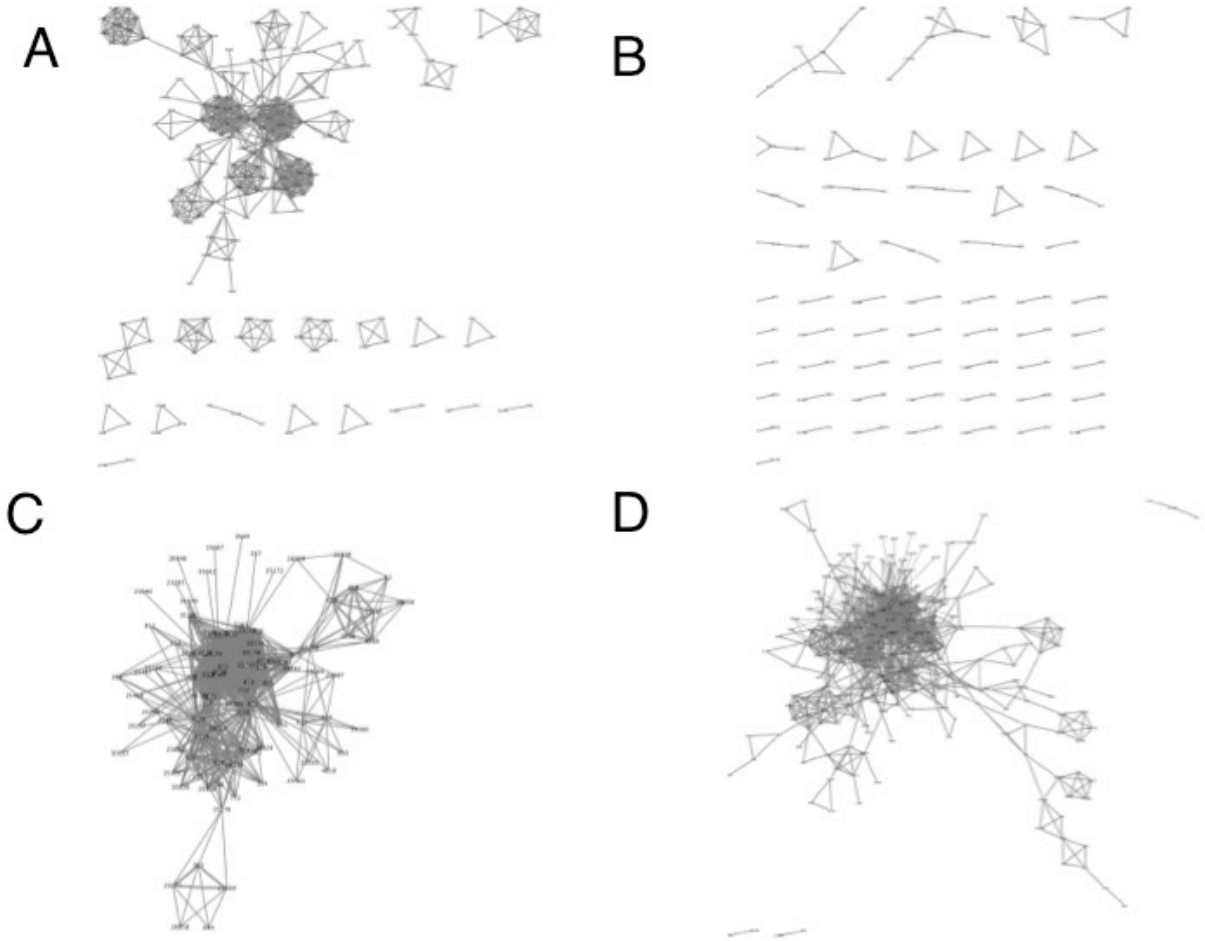


**Figure A2.** The connected components in the networks of coauthorship (A), common institutional affiliations (B), common media outlets (C), and in the union of all three (D), for year 2012.

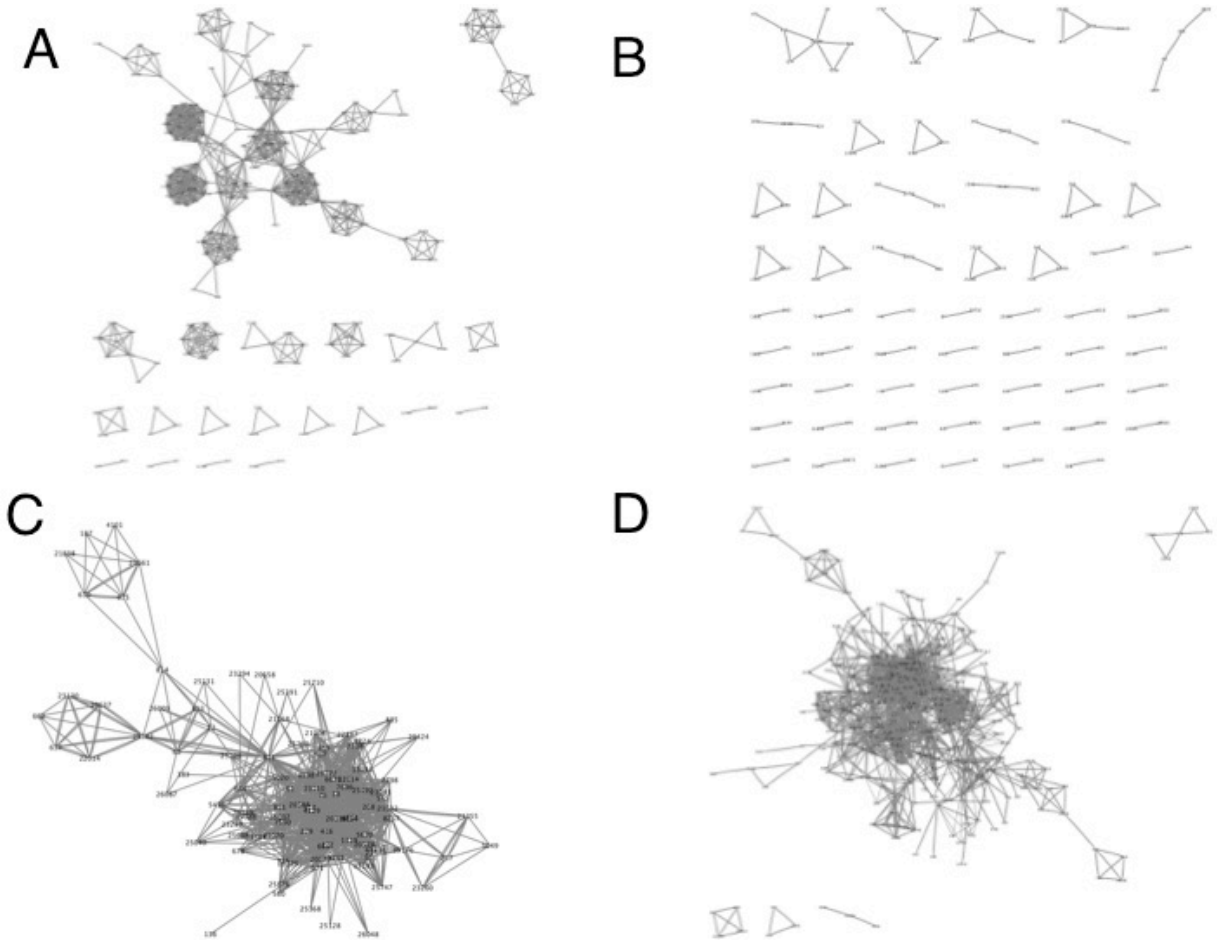


**Figure A3.** The connected components in the networks of coauthorship (A), common institutional affiliations (B), common media outlets (C), and in the union of all three (D), for year 2013.

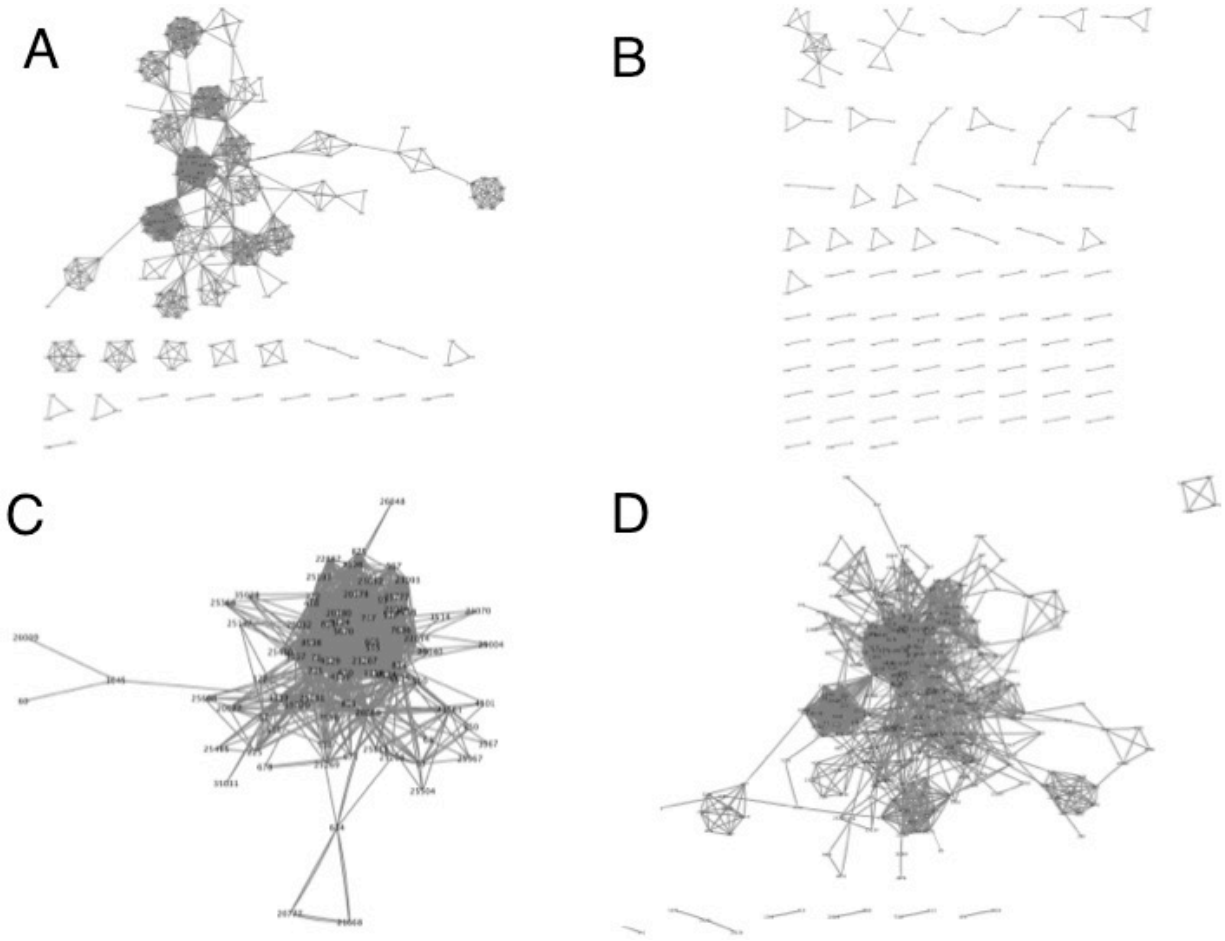


**Figure A4.** The connected components in the networks of coauthorship (A), common institutional affiliations (B), common media outlets (C), and in the union of all three (D), for year 2014.

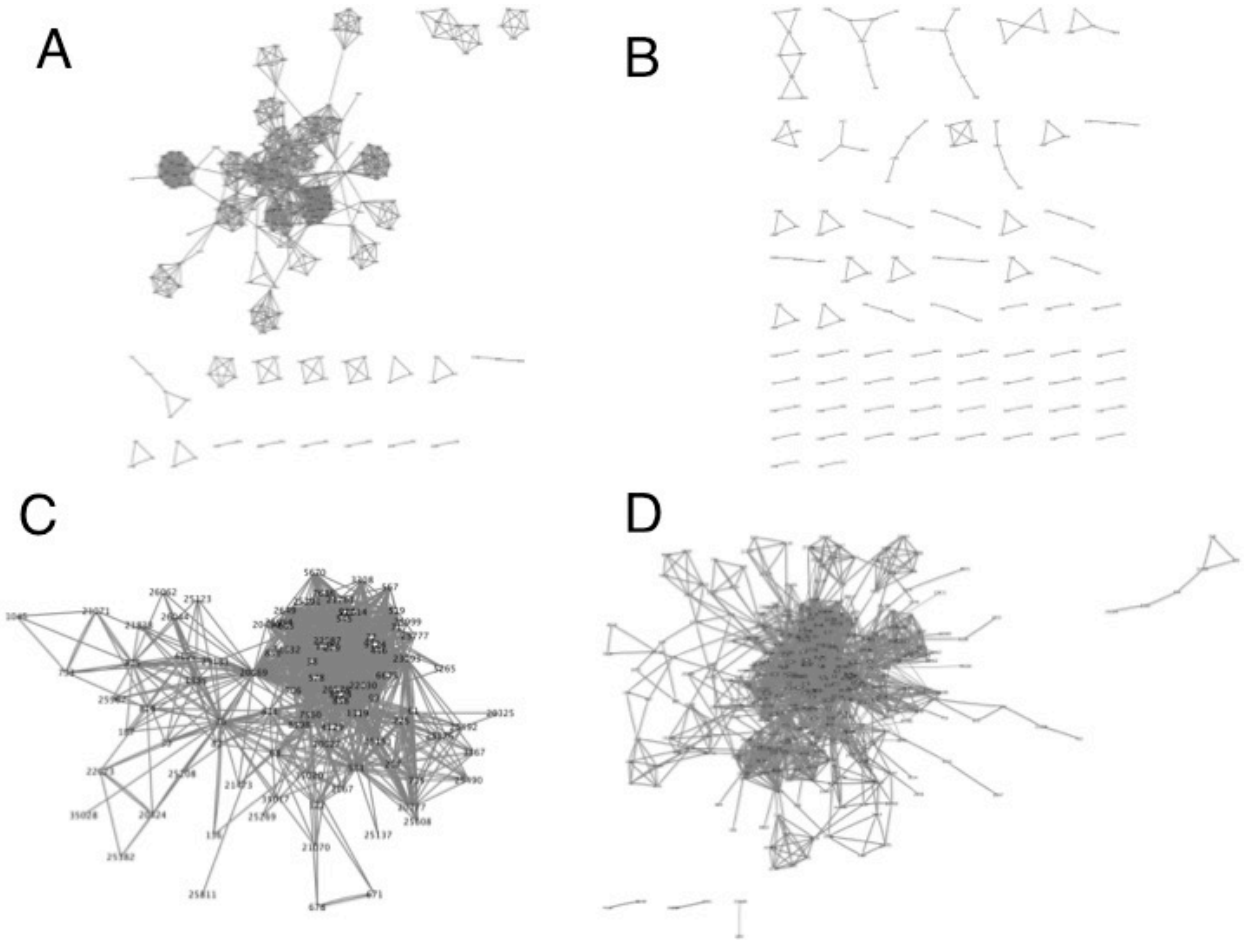
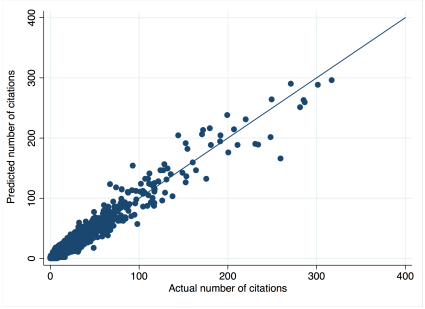
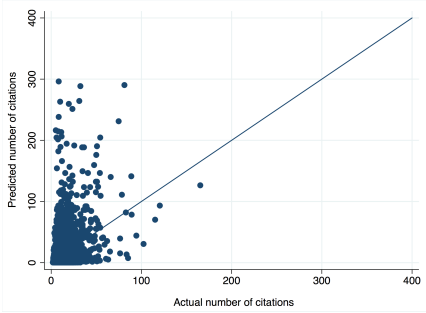


**Figure A5.** Models fit: Predicted vs. actual number of citations, by model and centrality measure.

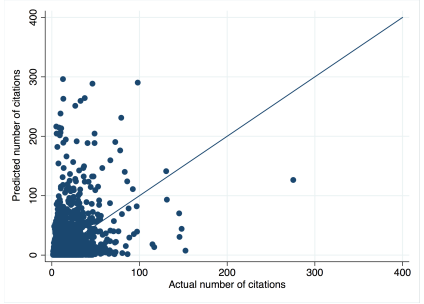
Conditional eff. Poisson model, degree centrality as dep. var. (Tab. 2, col. 3)



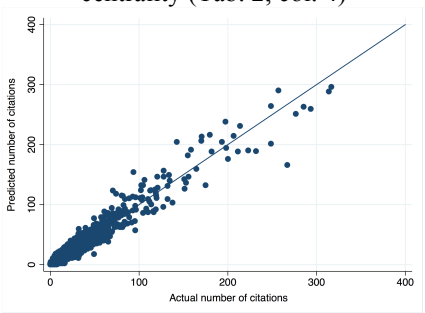

Pooled PPML model, degree centrality as dep. var. (Tab. A4, col. 7)



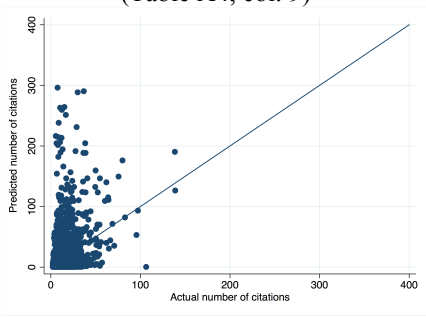

Negative binomial model, degree centrality as dep. var. (Tab. A5, col. 7)



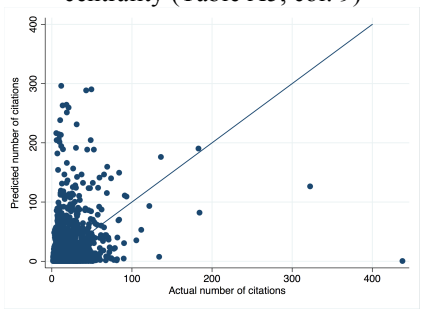

Conditional eff. Poisson, betweenness centrality (Tab. 2, col. 4)



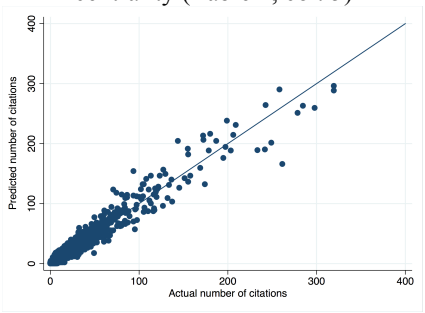

Pooled PPML, closeness betweenness (Table A4, col. 9)



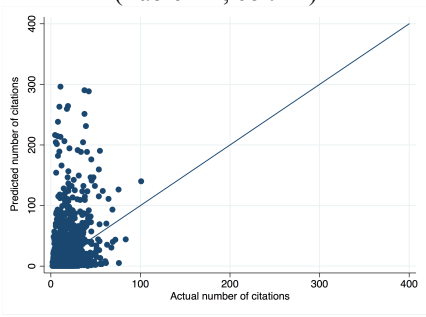

Negative binomial, betweenness centrality (Table A5, col. 9)




Conditional effects Poisson, closeness centrality (Table 2, col. 5)




Pooled PPML, closeness centrality (Table A4, col. 11)




Negative binomial, closeness centrality (Table A5, col. 11)