The strength of strong ties: How co-authorship affect productivity of academic economists?

Giulio Cainelli, Mario A. Maggioni, T. Erika Uberti & Annunziata de Felice

Volume 102 · Number 1 · January

2
0
1
5

FOUNDED IN 1978

Scientometrics

An International Journal for All
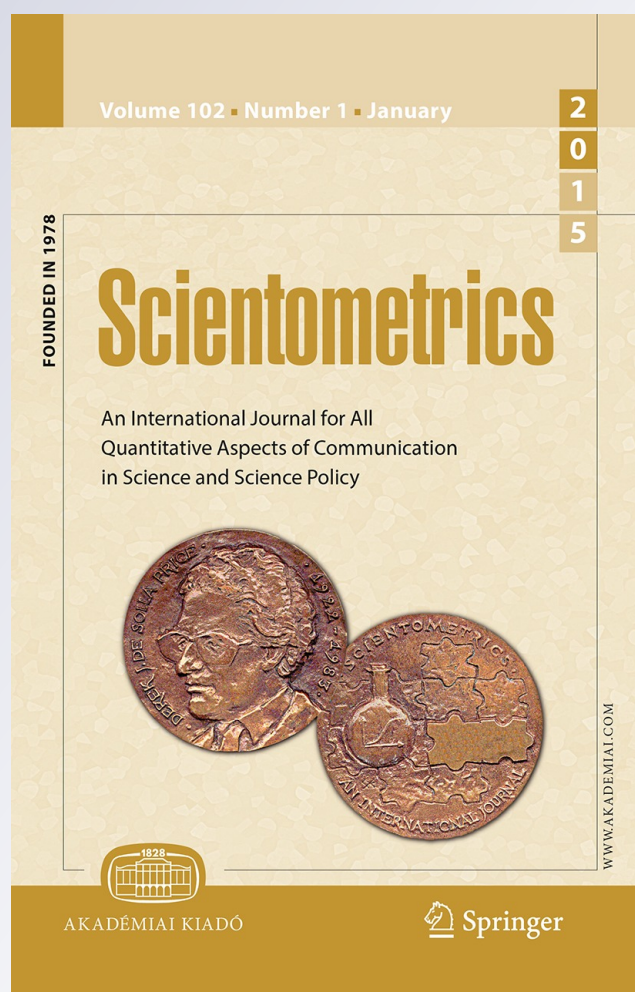Quantitative Aspects of Communication
in Science and Science Policy

WWW.AKADEMIAI.COM

1828
AKADÉMIAI KIADÓ

Springer

Springer

Springer

The strength of strong ties: How co-authorship affect productivity of academic economists?

Giulio Cainelli · Mario A. Maggioni · T. Erika Uberti · Annunziata de Felice

Abstract   Increased specialization and extensive collaboration are common behaviours in the scientific community, as well as the evaluation of scientific research based on bibliometric indicators. This paper aims to analyse the effect of collaboration (co-authorship) on the scientific output of Italian economists. We use social network analysis to investigate the structure of co-authorship, and econometric analysis to explain the productivity of individual Italian economists, in terms of 'attributional' variables (such as age, gender, academic position, tenure, scientific sub-discipline, geographical location), 'relational' variables (such as propensity to cooperate and the stability of cooperation patterns) and 'positional' variables (such as betweenness and closeness centrality indexes and clustering coefficients).

The Econ tribe occupies a vast territory in the far North. Their land appears bleak and dismal to the outsider, and travelling through it makes for rough sledding; but the Econ, through a long period of adaptation, have learned to wrest a living of sorts from it. (…) The extreme clannishness, not to say xenophobia, of the Econ makes life among them difficult and perhaps even somewhat dangerous for the outsider. This probably accounts for the fact that the Econ have so far not been systematically studied.

G. Cainelli
Dipartimento di Scienze Economiche 'Marco Fanno', Università degli Studi di Padova, Via del Santo 33, 35123 Padova, Italy

M. A. Maggioni · T. E. Uberti (✉)
DISEIS and CSCC, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milan, Italy
e-mail: erika.uberti@unicatt.it

A. de Felice
Dipartimento di Giurisprudenza, Università degli Studi di Bari "Aldo Moro", piazza C. Battisti 1, 70121 Bari, Italy

Information about their social structure and ways of life is fragmentary and not well validated. More research on this interesting tribe is badly needed.
Axel Lejonhufvud (1973, p. 327) our italics

Introduction

Co-authorship is an increasing phenomenon in science and economics makes no exception to this trend. Laband and Tollison (2000) show that in 'three prominent economics journals: the *American Economic Review*, the *Journal of Political Economy*, and the *Quarterly Journal of Economics*', the percentage of co-authored papers grew steadily between 1950 and 1994, <10 % in the 1950s and to around 70 % in 1994. Their data perhaps overemphasize the phenomenon, but are also consistent with the results of other studies. For instance, McDowell and Melvin (1983), for a sample of eight major economic journals, found the percentage of co-authored papers was about 30 % in 1976; Hudson (1996), based on a similar sample, found that the average number of co-authored papers in 1993 was over 50 %; Sutter and Kocher (2004) analyse 15 economics journals with the highest average impact factors, for the period 1977–1997, and found 44 % of papers on average were co-authored; Medoff (2003) surveyed 'thirty-one top economics journals', and found the percentage of papers with two authors, in 1990, was approximately 40 %; and Vieira (2008), for a sample of 168 journals available in the 'economics' class in the ISI database, found 47 % of papers were multi-authored (two or more authors) between 1986 and 1996.

In a previous paper (Cainelli et al. 2012), thanks to an purpose-built original database, we describe the structure of the Italian academic economists community and analyse the positive effects of collaborative national and international behaviours on the individual productivity. In this paper we build on the same database to investigate the attributional (age, gender, academic position, tenure, scientific sub-discipline, geographical location), relational (propensity to cooperate and stability of cooperation patterns) and positional (betweenness/closeness centrality indexes and clustering coefficient of the co-authorship network) variables in determining the productivity of Italian economists. To achieve this aim, in section "Co-authorship: causes and consequences" we briefly review the existing knowledge on the determinants and effects of co-authorship in the social sciences; section "The original database" describes in details the original database used in our analysis. Section "Relational structure of co-authorships" introduces social network analysis (SNA) methodologies and describes a number of relational and positional indexes used in the subsequent econometric analysis. Section "The econometric analysis" presents the results of three econometric exercises to show the role of attributional, relational and positional variables in determining the productivity of Italian economists. Section "Conclusions" concludes the paper.

Co-authorship: causes and consequences

Recent research studies (such as: Kalaitzidakis et al. 1999; Goyal 2005; van der Leij and Goyal 2011) and policy reports show an increasing relevance of multiple authorship in almost every field of science. In the meantime, universities', government bodies' and funding agencies' throughout the world are evaluating scientific research output through 'objective' bibliometric indicators. These phenomena are strictly intertwined since the pressure on academics to publish has promoted more co-authorships based on the belief

that co-authorship increases both the quantity[1] and quality[2] of research output in the form of published papers (Lee and Bozeman 2005). Thus, as the 'publish or perish' policy diffuses across national and disciplinary borders, the appeal of co-authorship increases among researchers.

This section is thus devoted to the review of a substantial, but not very prominent stream of literature, devoted to the theoretical and, above all, empirical analysis of the determinants and consequences of co-authorship.

While most of the theoretical literature (see among others: Laband 1987; Laband and Piette 1995; Laband and Tollison 2006; Butler 2007) concentrates on complementarities in skills and attitudes, Fafchamps et al. (2010, 2006, p. 8) underline that 'if research output depends only on ability, collaboration is most likely between authors of a similar level of ability (assortative matching) but with non-overlapping competences (complementarity in competences)'. However 'collaboration between high and low ability authors can arise if the low ability author provides more effort. In this manner the time-constrained high ability author can produce more research while the low ability researcher produces better quality output' (ibid). In this process, each author faces a 'matching' problem in finding his co-author, and this problem is exacerbated by the fact that one of a co-author's qualities is based on an ex ante evaluation of an unobservable variable (his/her effort). However, since 'collaborating with someone reveals valuable information about their ability and motivation, it follows that a referral about a researcher *i* is particularly informative when it is provided by a previous coauthor of *i*. Referral by a coauthor can thus be construed as a vetting process, stating whether a coauthor is competent and can be trusted to do his or her share of the work' (ibid., p. 12).[3]

The scale and scope of the empirical literature is definitely wider and calls for a brief survey of the different determinants and facilitators of co-authoring highlighted over 20 years of research.

- Specialization: 'the explosion of knowledge in economics, as the sheer growth of knowledge resulted in increased efficiencies of specialization and co-authorship relative to working alone' (McDowell and Melvin 1983, p. 156).

  This argument has been developed by several other authors:

- Multi-disciplinarity: since it is often fruitful to bring different perspectives to the study of a single issue, this is easily achieved by a 'multi-disciplinary configuration of the research

---

[1] The empirical evidence for both arguments is mixed. In terms of quantity, McDowell and Smith (1992) use cross-sectional data on academics and regress the number of articles produced by an individual (with co-authored articles discounted by the number of authors) on the percentage of co-authored articles and find no significant result. Durden and Perri (1995) use time series data for annual economics publications, over 24 years, and find that the total number of publications is positively related to the number of co-authored publications. Hollis (2001) uses the same data, and regresses total publications on the *proportion* that is co-authored, and finds no significant relationship. He then employs panel data for 339 individuals (US and Canadian AEE members) and finds that, if publications are discounted by number of authors, 'adding one more author is associated with a per capita reduction in output of between 7% and 20%' (Hollis 2001, p. 527).

[2] In terms of quality also, the empirical evidence is mixed. Laband and Tollison (2000) document an acceptance rate of 12% for collaborative papers submitted during the mid-1980s to the *Journal of Political Economy*, compared to 10% for single-authored papers. Others, such as Laband (1987), measure quality by citation frequency and report that this index is significantly higher for co-authored articles, while Barnett et al. (1988) use the positioning of the article in the journal as the signal of quality and find no support for this argument.

[3] These cites theoretically support the inclusion of a series of network analysis indexes in the empirical part of this paper, to take account of both the attributional features of each author and his/her relational and positional characteristics.

team'[4] (Sigelman 2009, p. 508). Further: 'As authors work in areas outside of their major areas of specialty they tend to engage in co-authorship to a larger extent than the authors who published within presumably more familiar areas in terms of interest or expertise' (Piette and Ross 1992, p. 281).Technological complementarities: the need to master very different analytical tools increases the importance of different expertise in order to manage numeric simulation packages, econometric estimation applications, and huge data bases[5] (Hudson 1996). This point is further re-enforced as follows: 'Only one specific type of complexity that matters. The combination of quantitative methods with other economic analyses is what appears to spurn co-authorship' (Nowell and Grijalva 2011).

- Synergy: the gains from collaborative work may be the result of a sort of synergy where 'multiple contributors develop ideas that none would have developed on his or her own. Synergy differs from skill complimentarily in the sense that it can exist between individuals with very similar skill sets' (Hudson 1996, p. 157). Johnston et al. (2013) highlight that citations patterns in economics change over time and are different for empirical and theoretical papers. Regarding the co-authorships issue, they underline that the co-authorship is positively correlated with citation process since it indicates 'the transfer of knowledge between those who created the original idea' (Johnston et al. 2013, p. 1024).

- Opportunity cost of time: the increased emphasis on research output and the use of publication output as a criterion for promotion, increases the opportunity cost of time for typical researchers in economics. Every other activity than doing research and writing papers, tends to be shelved or minimized. This means that acknowledgement is no longer a sufficient 'reward' for pre-submission review by a colleague in the field and 'this increased "price" often takes the form of co-authorship' (Barnett et al. 1988, p. 540).

- Risk diversification: 'the editorial review process contains a random element that many would agree is large (…) and a given review may motivate a rejection, revision, or acceptance depending upon the editor's judgement concerning reader interest, the size of the journal's backlog, or a host of other potential factors (…). Thus, the author of a paper faces considerable uncertainty (…) A natural response is to diversify against this risk by co-authoring papers. Through co-authorship, one is able to increase the total number of papers submitted within a given period of time, thereby reducing the variance of the random element inherent in the review process. Thus, even if the value of co-authored papers is discounted exactly by the number of authors, and if there are no synergistic or quality effects in co-authoring, there will still be incentives to collaborate' (Barnett et al. 1988, p. 540).

- Assigned value of co-authored papers: department 'chairmen ordinarily "assign a weight" to coauthored papers that exceeds l/n (with n being the number of authors), presumably to encourage collaborative research' (Liebowitz and Palmer 1983, quoted in Sauer 1988, p. 857). Many universities and funding agencies promote and reward collaboration in the belief that it has a positive impact on research productivity (Laband and Tollison 2000).

- Social interactions and pressures: collaboration may be chosen based on consumption/leisure reasons. Working with co-authors 'offers opportunities for friendship and camaraderie' and is a way to escape academic isolation (Medoff 2003, p. 607; Acedo et al. 2006; Holder et al. 2000). Having co-authors acts also as a motivation to keep to self-imposed deadlines.

---

[4] 'Interdisciplinary research also should be characterized by high rates of co-authorship by the same reasoning. The Piette–Ross insight, coupled with economists' steadily increasing colonization of other disciplines during the latter half of the twentieth century, may ex-plain a significant portion of the increase in the incidence of co-authorship' (Laband and Tollison 2000, p. 640).

[5] 'It may be cheaper for an individual to acquire new capacity (human capital) to produce through formal collaboration (merger) with someone who already has the requisite human capital than to acquire the needed knowledge de novo, personally' (Laband and Tollison 2000, pp. 639–640).

We can add also that innovations in information and communication technology (ICT) have made collaboration, even at a distance, easier and less costly[6] although Sutter and Kocher (2004) based on a gravity model of the co-authorship patterns in US departments, and based on the top 15 economics journals, find that the coefficient of the variable for geographical distance is non-significant for all the periods investigated (1977, 1982, 1987, 1992, 1997).[7]

Finally, it must be acknowledged that co-authorship has some negative effects, otherwise every paper would be written in collaboration, and the number of collaborators per paper would be infinite. Looking at the 'dark side' of co-authorship, we may mention the following reasons acting against this collaborative practice:

- Compromise: 'An individual author working with a group will have to agree to a certain approach, certain text, even certain conclusions that that person might not enunciate in the same way if working alone. Because multiple authorship inevitably involves compromise, thus it tends to reduce risk taking in academic papers. The result may be more technically proficient papers than in the past, but at the cost of the imaginative leap forward that starts economics in a new direction or gives fresh impetus to an old subject area' (Hudson 1996, p. 157).
- Organization and communication costs: 'multi-authored papers impose costs of organization and communication that may lead to diseconomies of scale. These are probably greater if all the collaborators are equally involved with all parts of the research and all parts of the paper. Developments in technology in recent years may have reduced the threshold at which these problems occur, but at some level they surely continue to exist' (ibid., pp. 157–158).
- Reward structure: 'Any net advantage of collaboration may disappear altogether if some individuals combine even though the sum of what each could achieve working alone exceeds their combined efforts. This may occur if an economist can achieve a greater gain in academic reputation from multi-authored rather than single-authored papers' (ibid., p. 158). This results is confirmed by Hilmer and Hilmer (2005) who find that the estimated return (as measured by increases in annual salaries) to a sole authored article is twice the estimated return to an article with more than one author (Hilmer and Hilmer 2005, p. 520).

From the above, it is evident that we need to consider both 'positional' and 'relational' variables in estimating the determinants of scientists' productivity. To achieve this, we built an original database in order to calculate positional and relational indexes for use as regressors in our econometric exercise.

The original database

Our analysis is based on an original dataset built by matching two different data sources: (1) the Italian economists population drawn from the official database of the Italian

---

[6] 'The development of technology has made collaboration more accessible across time and space. Overnight mail, photocopiers, computers, fax machines, email, and teleconferencing make long-distance collaboration considerably less daunting and time consuming. In essence, the invisible college of the 1960 s and 1970 s has been replaced by the "virtual college," or, more appropriately, the "virtual research center," of today' (Fisher et al. 1998, pp. 847–848; Kretschmer 1994). In order to detect the relevance of networks in publishing, an interesting perspective on the analysis of the editorial boards has been conducted by Baccini and Barabesi (2010).

[7] As for the effects of ICT on co-authorship patterns, Butler (2007) shows that (at least for a subsample of American Economic Association members working in the JEL fields D8, G2 and J3) the rate of internet penetration moved from almost nothing in 1995 to around 60 % in 1997, and to almost 100 % in 1999; while Maggioni et al. (2009) show that the average distance between co-authors working on the issue of 'industrial clusters', increased continuously in the period 1969–2007.

Ministry of Universities and Research, managed by Cineca (henceforth MIUR-Cineca) and (2) their scientific production extracted from the Econlit database of the American Economic Association.

The entire population of Italian academic economists, at 31st December 2006, is composed by 1,620 individuals holding one of the following academic positions: Tenured Full Professor (TFP), Full Professor (FP), Tenured Associate Professor (TAP), Associate Professor (AP), Senior Lecturer (SL) and Lecturer (L).[8] According to MIUR,[9] economists are sub-divided into six disciplinary groups; however for the empirical analysis we re-classified these groups into four sub-disciplines: *Economics*, *Econometrics*, *Public Economics* and *Others*.[10] Table 1 presents the distribution of academic positions and scientific fields in the Italian academic population of economists.

Once the population of Italian economists has been identified as described above, their scientific production can be traced in the Econlit database[11] recorded for the period 1969–2006. These records were downloaded between August 2007 and February 2008, and painstakingly corrected for errors in people's names and double entries. Since the aim of this paper is to analyse the effects of co-authorship on scientific productivity of Italian economists, we chose to measure the productivity of an individual as the total number of journal articles (henceforth: JA) published, per year.[12] The total number of JA published by the Italian economists in the period under analysis is equal to 8,679.

Before describing in details the Econlit database, it is useful to highlight its strengths and limitations. Dolado et al. (2003) suggest that Econlit contains missing information on authors and includes some errors and omissions with respect to publications, number of pages, etc. Furthermore, especially in the 1970s, the geographical coverage was limited to certain countries and to international publishers. Finally, the database does not include any evaluation or publication weight (i.e. impact factor or similar) for "scientific product".[13] However, the use of Econlit allowed to include, in the analysis, the publication activity on Italian scientific journals which are almost absent in the ISI-Thomson dataset and to record

---

[8] These positions in Italian are: *Professore Ordinario, Professore Straordinario, Professore Associato Confermato, Professore Associato, Ricercatore Confermato, Ricercatore*. There is another position, *Assistente Ordinario*, which is between a lecturer and a professor whose academic duties are similar to a senior lecturer. Since this position is increasingly disappearing (percentage is only around 1%), we code this as SL. See Appendix A1 and Cainelli et al. (2006) for more details.

[9] The law is contained in the *Decreto Ministeriale 4 ottobre 2000* and published in the *G.U. n. 249 del 24 ottobre 2000—supplemento ordinario 175*.

[10] *Economics* also includes Political Economy and Economic Policy (i.e. corresponding to the disciplinary sectors SECS-P/01 *Economia Politica*, SECS P/02 *Politica Economica*), *Econometrics* refers to the disciplinary sector SECS-P/05 Econometria, *Public Economics* to SECS-P/03 *Scienza delle Finanze*, and *Others* is a miscellaneous disciplinary sector which includes SECS-P/04 *Storia del Pensiero Economico* (History of Economic Thought), and SECS-P/06 *Economia Applicata*, a mix of regional economics, transport economics, and industrial organization.

[11] Econlit distinguishes between different "scientific products": journal articles, collective volume articles, books, working papers and dissertations.

[12] The focus on JA reflects the overwhelming role played by these "products" in the recent evaluation procedures (CIVR, VQR) put forward by individual universities and the Ministry.

[13] By using Econlit as bibliographic database, we focus at the quantitative profile of the internationally-visible scientific production of Italian economists, but we do not measure the "scientific value" of their publications. Although this might be considered a significant shortcoming, we do not believe to be as serious as the dominant faction in the current debate on evaluation might suggest. We are in fact convinced that the issue of how to weigh publications, and particularly the use of impact factors or citation indexes, should still be considered an open question for economics, as suggested by the results of the evaluation of European economics departments carried out by the *European Economic Association* (see Neary et al. 2003a, b).

Table 1 Academic position and scientific field

Scientific field	L	SL*	AP	TAP*	FP	TFP*	Total
Economics	13.9	15.0	10.6	18.5	8.0	34.0	1,153
Econometrics	13.1	13.1	9.8	13.1	13.1	37.7	61
Public economics	13.5	16.0	4.5	14.0	8.0	44.0	200
Others	8.7	13.6	14.6	22.8	10.2	30.1	206
Total	214	241	167	296	137	565	1,620

Source: our calculations based on MIUR-Cineca database at 31 December 2006

Percentage (by rows) and total number at 31 December 2006

* These academic positions are tenured

data on collective volume articles (CVA) which have been used as instruments in the econometric analysis (see section "The econometric analysis").

Relational structure of co-authorships

This section introduces the SNA indexes and analytical tools, and their application to study the structure and evolution of the scientific collaboration behaviour of Italian economists.

Selection and treatment of networks

SNA is a scientific method of analysis that investigates the structure of the relations between social units of analysis, using graph theory, mathematics and statistics tools. Key SNA concepts are *actors* (i.e. players or nodes in the *network*); *relational ties* (i.e. the links connecting the actors) and *groups/subgroups* (i.e. the subsets of the actors and the relations among them).

Having defined the sample of actors and relations to be studied, SNA distinguishes the type or *mode* of the networks, i.e. the 'number of sets of entities on which structural variables are measured' (Wasserman and Faust 1994, p. 35). Most SNA studies relate to 'one-mode networks' (i.e. where all entities or nodes belong to the same set of actors); however, some deal with *two* sets of social entities[14] (and the relations connecting one set to the other). These are called 'two-mode networks' or 'affiliation networks'.

In this paper we build a two-mode network, where one set of nodes, i.e. mode 1, is composed of journal articles (JA) and the other set of nodes, i.e. mode 2, is the papers' authors (AU) (see Fig. 1a). We transform this into a 'one-mode network' in which AU are the nodes and the papers are the links between co-authors (see Fig. 1b).

More formally Fig. 1a is a two-mode network defined as $\mathbf{X} = (JA, AU, R, w)$, where the two disjoint sets are respectively JA, i.e. 8,679 articles, and AU, i.e. 2,972 authors included in the analysis, the scientific collaborations are denoted as $R \subseteq JA \times AU$ and the mapping $w{:}R \rightarrow R$ represents a weight, i.e. the number of co-authored articles. The resulting affiliation matrix, $\mathbf{X}$, is rectangular, i.e. 8,679 JA times 2,972 AU, with links originating from different authors and targeting the same paper representing cases of co-authorship. In order to emphasize the structure of co-authorship among economists we transformed the
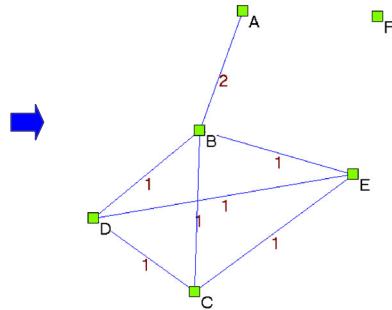
---

[14] These sets can include 'actors' and 'social events', i.e. members and administrative boards, authors and papers, citations and patents, etc. (Wasserman and Faust 1994; Borgatti and Everett 1997; Doreian et al. 2004).

Mode 1 – Nodes = Journal Articles: 1, 2, 3, 4, 5…
8,679[th]
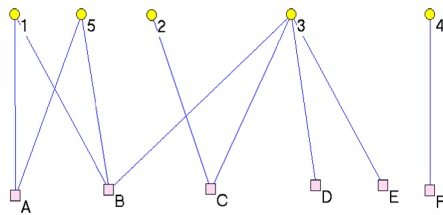
Mode 2–Nodes= Authors: a, b, c, d, e, f … 2,972[nd]

Links connects authors to papers (i.e. paper 1 has been written by authors, A and B)

Nodes = 2,972 Authors;

Links = 8,679 co-authored papers (i.e. the link between A and B is valued 2 since A and B have co-authored 2 papers)

(a) two-mode network   **(b)** one-mode network

Fig. 1 One-mode versus two-mode representation of co-authorship networks

two-mode network into a one-mode network (depicted in Fig. 1b) where the network is defined as $\mathbf{X_1} = (AU, R_1, w_1)$, representing the collaborations, R_1, among authors AU, i.e. a 2,972 times 2,972 squared matrix where nodes are authors, and w_1 represents the number of scientific collaborations between dyads of authors. Then we performed a series of SNA analyses on the one-mode network for the whole period 1969–2006 and—given that co-authorship changes radically over time—we investigated the networks for four different historical periods (1969–1976; 1977–1986; 1987–1996 and 1997–2006).

Before discussing the structural features of collaborative networks, we describe how the network of 2,972 authors is defined. Figure 2 synthesizes the final population of economists investigated, defined using the snowball sampling procedure, from 1,620 Italian economists selected from the MIUR-Cineca database.

First, we identify the entire population of Italian economists according to the MIUR-Cineca database (**M**) as defined in section "The original database", i.e. the 1,620 Italian economists[15] in an official academic position at the end of 2006. For each individual we identify the records indexed in the Econlit database, including details of year of entry, full records on JA, CVA, books and dissertations. This snowball procedure means that the number of individuals involved in writing JA increases with respect to the initial population **M** and we can identify a new Econlit population, **E**, that includes individuals belonging to set **M**, and all their co-authors (if any) including each affiliation,[16] for a total of 2,972 individuals.

The intersection between sets **M** and **E** produces the intersection set **P**, which is composed of 1,317 'Italian' Academic Economists who wrote at least one JA indexed in Econlit[17] and 3 other complementary sub-sets: **N**, **O** and **F**. Subset **N** is composed of 262

---

[15] The 19 foreign economists affiliated to Italian universities in the economics fields, in this paper are considered to be 'Italian economists'.

[16] The affiliations can be worldwide: this selection includes individuals with both foreign and Italian affiliations (i.e. individuals from other scientific sectors not included in those defined in "The original database", or institutions outside academia, i.e. Banca d'Italia, CNR, ISTAT, etc.).

[17] We should stress that 41 individuals in the **M** population have entries in the Econlit database that do not refer to JA.

Fig. 2 Set representation of the populations in the dataset
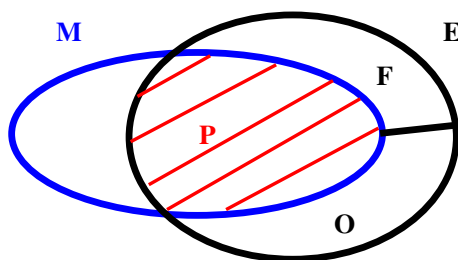
'Italian' academic economists with no JA entries in Econlit; **F** is composed of 806 'foreign' economists with at least one co-authored JA recorded in Econlit, involving one or more Italian economists (included in **P**); **O** is composed of 849 'other Italian' co-authors with the economists in set **E**, which do not belong to **P** because: (1) they are affiliated to non university institutions (e.g. Bank of Italy, CNR-National Research Council, ISTAT-National Statistical Office, foreign and international institutions); or (2) they belong to non economic sub-disciplines (i.e. business management, statistics, etc.).[18]

This 'partial snowballing' sampling procedure allows us to identify the structure of the scientific collaboration of the Italian academic economists (in set **P**). Since in this article we are interested in the Italian economists we do not investigate in detail the structure for the residual **F** and **O** populations.[19]

SNA indexes and network topology

Since our interest is in co-authorships and the effects on scientific production, we treat the data on the publications of Italian economists transforming the two-mode network into a one-mode network as described previously, and we calculate a number of different indexes measuring the structure of the network, the positions and relational roles of each node to be used as inputs for the econometric analysis in section "The econometric analysis".

In order to synthesize the general features of the overall network, we calculate its density values (d and d_r), degree centralization (C_{deg}), clustering coefficient (CC), average path length (APL), diameter (δ), and average degree (*av_deg*).

Density (d) is defined analytically as the ratio between the total number of actual links (L) and all possible potential links among all nodes in the network (n):

$$d = \frac{2L}{n \times (n-1)} \tag{1}$$

This network index ranges from 0 (i.e. the network is disconnected) to 1 (i.e. the network is complete; i.e. all possible links are present) and represents the completeness of the network (Wasserman and Faust 1994).

---

[18] Since we are interested in investigating differences in the collaborative behaviour of Italian economists in scientific collaborations with native academics and collaborations with foreigners, and since researchers are often quite mobile (and may change affiliations in the course of their careers), we attribute the residual population to sub-sets **F** or **O** based on nationality.

[19] The size and percentage coverage of a network is an open issue in SNA, see Ter Wal and Boschma (2009) and Maggioni and Uberti (2011) for further details.

Since it is not possible to compare densities for networks of different dimensions, we follow Maggioni ([1995](#)) and compute a relative density[20] value (d_r):

$$d_r = \frac{L - (n-1)}{\frac{n \times (n-1)}{2} - (n-1)} \tag{2}$$

Degree centralization index ($C_{\deg}$) is a measure of the variance in the degree centrality values of the nodes in a given network (Freeman [1979](#)), expressed analytically as:

$$C_{\deg} = \frac{\sum_{i=1}^{n} \left[C_{\deg}(a^*) - C_{\deg}(a_i) \right]}{n^2 - 3n + 2} \tag{3}$$

where $C_{\deg}(a^*)$ and $C_{\deg}(a_i)$ respectively are degree centrality index (i.e. the number of a node's direct links) of the most central node, and degree centrality of a generic node i. As in the case of d, the index ranges from 0 (i.e. all nodes have the same degree centrality index) to 1 (i.e. there is one node that connects the entire network). This index, therefore, measures the hierarchization in the network and the presence (or absence) of pivotal node(s), i.e. a node(s) with a direct relation to most of the other nodes in the network.

While these network indexes occur frequently since early applications of SNA to sociological analysis, CC, APL and δ were recently introduced by mathematicians, physicists and computer scientists (Strogatz [2001](#); Albert and Barabasi [2002](#); Newman [2001](#), [2003](#)). These indexes (and the underlying degree distribution) are computed to classify large and complex structures (i.e. networks characterized by thousands of nodes) with reference to standard ideal types (e.g. random, regular, scale-free or small world structures).

The clustering coefficient of node i (CC_i) is the ratio of the existing links connecting a node's neighbours to each other to the maximum possible number of such links. Its average across all nodes in a network (CC) summarizes the extent to which the nodes in a graph tend to group together (Watts [1999](#)). More formally:

$$CC_i = \frac{\Lambda_i}{v_i} \tag{4}$$

and

$$CC = \frac{\sum_{i=1}^{n} \frac{\Lambda_i}{v_i}}{n} \tag{5}$$

where Λ_i is the number of edges in the neighbourhood of node i, v_i is the total number of possible edges of node i. CC_i and CC vary between 0 (i.e. no neighbour of any vertex is adjacent to any other neighbour of vertex i) and 1 (i.e. a node's neighbours are also neighbours of each other).

Average path length (APL) is a measure of the degree of separation between two nodes along the shortest path of their intermediaries, i.e. the average number of intermediaries between all pairs of actors in the network (Uzzi and Spiro [2005](#)). If APL is low, the actors in the network are close together and flows across the network are easy. Another measure of average distance in the network is diameter (δ), which

---

[20] For connected networks, the range is between 0 (i.e. the network is minimally connected, meaning that the removal of just 1 link would disconnect the whole network) and 1 (i.e. the network is complete, meaning all possible links already exist).

measures network connectedness and corresponds to the maximal distance between any pair of nodes, i.e. the longest geodesic path, and therefore will be a positive number >1. In a disconnected network, δ is equal to infinity, but it can be computed for connected components.

To classify the topological structure of a network it is usual to compare the indexes of the 'actual' or 'real' networks, to equivalent (i.e. with the same number of nodes and edges as the real network) random networks (Watts 1999; Uzzi and Spiro 2005).[21,22] A 'small world' network is a network in which most nodes are not neighbours of one another, but are separate only by a small number of steps.[23] Watts (1999) synthesizes some of the structural features of the topology comparing a given network's APL and CC with the same indexes calculated for an equivalent random network.

In this paper we check whether the network of Italian economists is structured according to a 'small world' typology by computing the 'small world quotient' (Uzzi and Spiro 2005), Q_{sw}, as follows:

$$Q_{sw} = \frac{\frac{CC_a}{CC_r}}{\frac{APL_a}{APL_r}} \tag{6}$$

where subscripts a and r respectively indicate the *actual* and the equivalent *random* networks CC and APL, in terms of average degree and density. The greater the Q_{sw}, (and particularly if the quotient is >1), the closer the structure of the network to a 'small world' structure.[24]

In order to investigate the evolution of the co-authorship network, we also study the relationship between network structure at time t and its evolution over subsequent periods. Albert and Barabasi (2002) find that the evolution of scale-free networks[25] generally follows a 'preferential attachment' process where central nodes become increasingly central because new nodes tend to establish proportionally more links with more central nodes.

In the following analysis we calculate $R(k)$, the relative probability of establishing new co-authorships, as follows:

$$R(k) = \frac{\text{proportion of new links to nodes of deg } k}{\text{proportion of nodes with deg } k_i} \tag{7}$$

If there is no preferential attachment, $R(k)$ is equal to 1, and describes a growth process in which new authors enter the co-authorship network by establishing links to existing

---

[21] It is useful to recall that a large random network (Erdős and Renyi 1959) is characterized by a binomial degree distribution, while CC—which depends on the size of the network—is equal to the ratio of average degree and number of nodes, and APL and δ depend on the size of the network structure and its average degree (Albert and Barabasi 2002).

[22] Similarly it is possible randomly to remove some nodes in order compare the main components of the real network with the equivalent random ones (Maggioni and Uberti 2009).

[23] In a 'small world' structure, the path between two nodes in very large network may be extremely short due to the existence of bridging agents. This means that even in very large and locally clustered networks efficient and fast information diffusion is possible.

[24] In the paper we compute Q_{sw} for the whole network (from 1969 to 2006) as well as for the last two sub-periods when a main component emerged in the network structure.

[25] These networks are characterized by a very skewed degree distribution with very few pivotal nodes, and a large number of peripheral nodes.

authors randomly. If $R(k)$ is >1, then growth follows a preferential attachment process because more central authors are more attractive to new co-authors.

The following network indexes have been computed for each node (author) and used in the empirical analysis: degree, betweenness and closeness centrality,[26] and clustering coefficient.

The degree centrality[27] index is a measure of the number of direct links connecting a given node, which in our case is the number of co-authors of each individual economist, which measures the local centrality of the economist.

Betweenness centrality is a measure of the number of times a vertex occurs on all geodesic (i.e. shortest) paths within a network connecting every node to every other node. This index identifies the strategic value of a node and its potential ability to control the relations of a network. Analytically:

$$C_b(a_i) = \sum_{j<k}^{n} \frac{g_{jk}(a_i)}{g_{jk}} \tag{8}$$

where $g_{jk}(a_i)/g_{jk}$ is the estimated probability of connecting two nodes and the numerator represents the number of geodesic distances connecting nodes j and k containing node i, while the denominator does not necessarily contain node i (Wasserman and Faust 1994).

Closeness centrality is an inverse function of the geodesic distances from one node to every other node in the network—hence it depends on both direct and indirect links—and reflects the efficiency of the communication channels of a given node with the rest of the network. Analytically:

$$C_c(a_i) = \left[\sum_{j=1}^{n} d(a_i, a_j) \right]^{-1} \tag{9}$$

where $d(a_i, a_j)$ is the number of lines in the geodesic path linking nodes i and j (Wasserman and Faust 1994).

Since these last indexes are computed for each node of the network (i.e. each Italian economist included in the analysis), we consider them regressors in the econometric analysis since they measure different aspects of the positional role of an author in the co-authorship network.

SNA of Italian economists

In order to detect whether and how co-authorship behaviour has changed over time, and whether the structure of co-authorship is similar to one of the ideal typical network topologies we analyse the co-authorship patterns of Italian economists using SNA techniques.

First, we observe that collaboration is evolving over time, and is becoming a more common phenomenon of Italian economists' scientific behaviour. For example, the number of co-authored JA has increased more than the number of single authored JA. In 1969, 20

---

[26] Degree, betweenness and closeness centralities are calculated as defined by Freeman (1979).

[27] On the basis of the degree centrality index of each node, we computed a network index, average degree (av_deg), as the mean value of the degrees (i.e. directly incident links) of each node in the network, in order to enable comparison of the collaborative behaviours of Italian academic economists in different time periods.

out of 24 articles recorded in Econlit were single authored; in 2003, 272 out of 540, while in 2006, 328 out of 567 articles were multi-authored.

The structure of the whole period network is very complex (see Fig. 3; Table 2). In a network of 2,972 economists there is a main component (MC), i.e. the largest group of connected nodes, which includes 2,061 economists, several of these wrote together more than one article (represented by the thicker lines between nodes), while the number of isolated nodes (i.e. economists always writing alone) is quite small, under 10 %. There are several sub-groups of different dimensions in the whole network, which represent small communities of economists (i.e. discipline-bounded or geographically delimited).

Table 2 shows the results for a network dichotomized according to a threshold value equal to 1. This procedure identifies collaborations for at least one JA, which means that multiple collaborations with the same co-author do not influence the structure of the network.

The density value shows that the network is quite sparse: density value is very low for both the whole network and the MC. The centralization measure indicates a fairly non-hierarchical network, with no stars in a pivotal role, in the whole network or in the MC. Diameter δ indicates that although the network is quite complex, the biggest distance between two connected economists is 30, which is a relatively small value.

The average degree values show that Italian economists collaborate with over two people in the whole network, and around three people in the MC.

In the complexity of Fig. 3 we identify nodes far apart and a main component. We should highlight that nodes far apart are not exclusively isolated (only 282 nodes are completely disconnected), but several other apparently isolated nodes (200) are diads, triads and sub-networks identifying people writing in economics sub-subjects. On the main component a typical collaborative pattern (defined also by the literature and detailed in the econometric analysis) emerges: some people collaborate more and this is captured by the thickness of the link.

There are some other interesting features that emerge when we compare different groups of economists, especially subset **P** with subsets **F** and **O,** the groups of foreign and 'other Italian' economists. In this work, the group of **P** economists represents 44 % of the whole network and 37 % of the MC.

If we consider the international openness of Italian economists in terms of the nationalities of co-authors, we find that foreign economists accounts for 38 % of the network and this percentage is 46 % for the MC. This suggests an overrepresentation of foreign economists in the most connected sub-community of co-authors.

Since cooperation behaviour is changing dramatically over time, we identify four distinct periods from 1969 onwards, which demonstrate how network features are changing (see Table 3 for basic SNA statistics).

In the first period (1969–1976) no particularly cohesive network structure can be identified, and the network is mostly disconnected. In the second period (1977–1986), the network structure is very disconnected, and we can identify several isolated nodes (i.e. economists writing alone). In both these networks the Italian economists community is not very open to collaboration with foreigners and is mainly representative of the selected scientific sectors. In both networks average degree is <1, and more co-authorship among the Italian economists community does not emerge until the late 1980s. Hence due to this pattern of isolated and diads (at maximum) MC were very unlikely to appear.

The structure of the networks changes radically in the last two periods: 1987–1996 and 1997–2006. A MC emerges, while the number of isolated authors remains fairly constant;
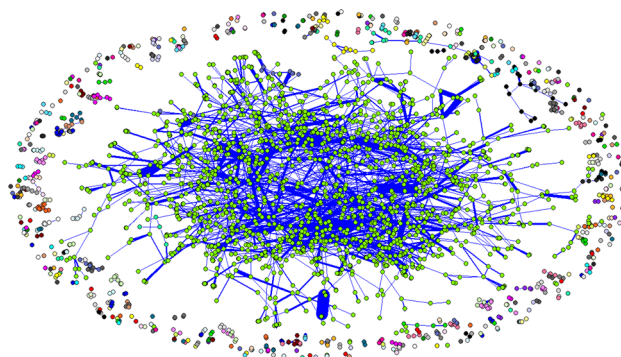
Fig. 3 The co-authorship network of Italian economists (1969–2006). *Note*: colours identify different sub-groups and the thickness of links identify the number of collaborations

Table 2 Basic SNA statistics of the whole network (1969–2006)

Index	Whole network	Main component
n	2,972	2,061
d	0.000824	0.001469
APL	8.292	8.296
δ	30	30
CC	0.569	0.552
C_{deg}	0.0103	0.0146
Isolated nodes	283	…
Isolated diads	113	…
Isolated triangles	43	…
Sub-networks with 4–9 nodes	40	…
Sub-networks with 11–15 nodes	4	…
Av_deg	2.447	3.026
SD	2.841	3.144
Min deg	0	1
Max deg	33	33

Note: Threshold value $= 1$

average degree increases and APL and diameter increase dramatically; and the maximum degree increased 6 times and a half, indicating a diffusive collaborative pattern. This shift in late 1980s could be due to the diffusion of the new ICT and to the intensification of post-doc experiences abroad.

The values for network density (both absolute and relative) are very low and are continuously decreasing over time.[28] This does not contradict the increasing values for average degree since, even if scientific collaborations increases over time, the increase in

---

[28] Relative density is negative, signalling the existence of isolated nodes.

Table 3 SNA statistics for different periods

Index	Period 1 (1969–76)	Period 2 (1977–86)	Period 3 (1987–96)		Period 4 (1997–2006)	
n	159	580	1,094		2,424	
size of MC	0	0		214		1,380
MC/n	–	–	19.56 %		56.93 %	
Size of P	126	445	692	89	1.176	570
Size of F	25	60	227	68	619	389
Size of O	8	75	175	57	629	421
d	0.0070	0.0014	0.0012	0.0120	0.0010	0.0022
d_r	−0.0056	−0.0021	−0.0006	0.0027	0.0001	0.0008
APL	1.328	1.751	8.180	8.533	8.282	8.295
δ	3	5	21	21	20	20
CC	0.458	0.606	0.535	0.531	0.596	0.564
C_{deg}	0.0221	0.0108	0.097	0.0447	0.0098	0.0167
Isolated nodes	92	293	295		263	…
n. isolated diads	34	55	88		114	…
n. isolated triads	5	16	43		46	…
Av_deg	0.553	0.786	1.364	2.561	2.328	3.041
Min deg	0	0	0	1	0	1
Max deg	4	7	12	12	26	26
SD	0.774	1.147	1.498	1.932	2.459	2.830

Threshold value = 1

these values is less than the increase in the number of possible links in a growing network.[29]

Table 4 presents the Q_{sw} index to detect the presence of a small world topology in the network structure of the Italian economists community. The value index (Q_{sw}) means that this network has features typical of a small world topology. Since the network is changing over time, and collaborative behaviour is increasing, we calculate this index for all the periods in which a MC is present.[30] All indexes are > 1, confirming the existence of a 'small world phenomenon' in the network of Italian economists.

Finally, to identify the forces behind the resulting topology of the co-authorship network we investigate the presence of preferential attachment. Figure 4 describes the existence of a non-linear dynamics of preferential attachment in the MC of Italian economists, passing from period 3 to period 4. As already detailed, a value >1 signals the existence of preferential attachment dynamics in the growth of the network, with more central nodes (i.e. node with a higher degree) in the previous period receiving a more than proportional share of new links in the subsequent period. This is the case for the Italian economists community in the two most recent decades. Authors with more than four different co-authors in the 3rd period tend to show an increase in the number of their co-authors. This

---

[29] We should remember that the denominator of density measures increases non-linearly ($n*n - 1$) with network size n.

[30] Thus we did not compute the index for periods 1 and 2.

Table 4 The 'Small World' of Italian economists

	Total (1969–2006)	3rd period (1987–1996)	4th period (1997–2006)
n	2,061	214	1,380
APL_a	8.296	8.533	8.295
CC_a	0.552	0.531	0.564
APL_r	6.739	5.349	6.383
CC_r	0.001	0.007	0.001
Q_{sw}	448.40	47.55	434.00
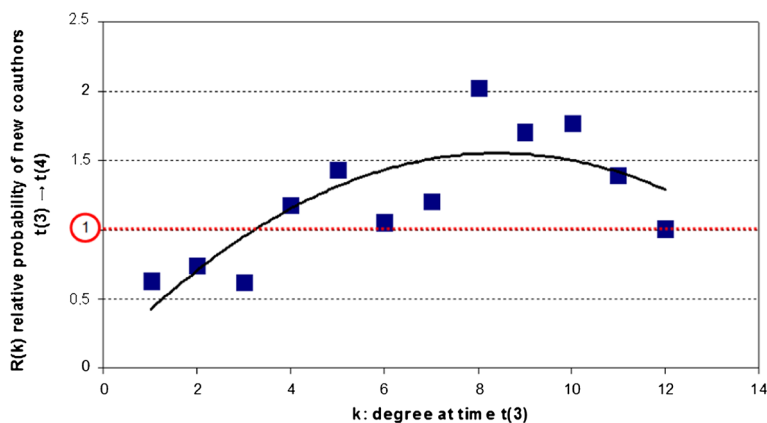
Fig. 4 Preferential attachment in the MC of the Italian economist network

may be due to different reasons, including that: (1) previous co-authorship signals both a willingness and ability of an author to cooperate with other scientists; (2) scientists tend to look for this kind of person when looking for potential co-authors.

The depiction in Fig. 4 suggests that the increasing returns to co-authorship are not linear; in particular, when the cumulated number of co-authors reaches 8, the relative probability of additional co-authors begins to decrease and becomes equal to 1 for 12 co-authors in period 3. This may be explained by an excessive level of transaction and communication costs for a subgroup of more than 12 co-authors, and a preference for stability in co-authorship relations, as described in section "The econometric analysis". We also calculate the distribution of degrees, and find that it is very skewed, and very different from a normal distribution, which confirms the wide variation in the collaborative behaviour of Italian economists and suggests that the network topology is more similar to a scale-free than to a random network.

The econometric analysis

We use different econometric methods to identify empirically the determinants of the probability to publish and of the scientific productivity of economists working in Italian universities. As stated in the introduction, and discussed extensively in section "Co-

authorship: causes and consequences", we expect that these determinants will have three distinct dimensions:

- an *attributive* dimension related to the individual characteristics of the researcher (measured by variables for gender, academic position, tenure, location);
- a *relational* dimension representing the researcher's connection to his/her scientific community (measured by variables such as the propensity for co-authoring, propensity to cooperate with foreign researcher);
- a *positional* dimension, which relates to the role of the researcher in the scientific community (measured by network variables such as betweeness centrality index and clustering coefficient).

In particular in Table 5 we identify the *attributional* determinants of the probability to publish using a probit model; in Table 6 we investigate the *relational* driving forces of an individual researcher's scientific productivity, through an Instrumental Variables (IV) estimation strategy; finally we focus on the impact of the *positional* characteristics of the researcher within the scientific community using OLS methods and analysing the sole MC of the co-authorship network.

While some econometric results on the attributional and relational determinants of both the probability to publish of an economist and his/her scientific productivity have been already presented in Cainelli et al. (2012), this paper focuses on the positional determinants by exploiting the SNA indexes computed in section "Relational structure of co-authorships" and using them as regressors in the econometric analysis.

The determinants of the probability to publish

To identify the *attributional* determinants of the probability to publish, we estimate a maximum-likelihood *probit* model as follows:[31]

$$Pr(publish_i = 1|X) = \Phi(X'_i\beta) \tag{10}$$

where Φ is the cumulative distribution function of the standard normal distribution, *publish_i* is a dummy variable that is equal to 1 if the economist has at least published one (column [1.] of Table 5), two, (column [2.] of Table 5) three (column [3.] of Table 5) or four (column [4.] of Table 5) journal articles (*JA*) in Econlit in the period 1969–2006, and 0 otherwise, and X indicates the regressors. These variables, constituted by a set of attributional dummy variables, are added stepwise in the econometric specifications, and are available for all the economists in our 2006 population. They include: (1) five geographic dummies (*North West*, *North East*, *Centre*,[32] *Islands* and *South*) describing the location of the economist's institution; (2) a dummy indicating whether the economist works in a Faculty of Economics and 0 otherwise (*Fac_economics*); (3) four dummies for economist's disciplinary groups (*Economics*, *Public Economics*, *Econometrics*, and *Others*[33]); (4) a dummy for gender (*Gender*) equals to 1 if the academic is male and 0 otherwise; (5) a dummy (*Tenured*) if the economist has a tenured academic position; (6) a dummy (NT*Lecturer*), that is equal to 1 if the economist is a non-tenured lecturer and 0

---

[31] Specifically, we use the *dprobit* model. Rather than reporting coefficients, *dprobit* reports the marginal effects, i.e. the change in the probability for an infinitesimal change in each independent, continuous variable and, by default, reports discrete change in the probability of the dummy variables.

[32] This has been used as the reference and not included in the regression.

[33] This has been used as the reference and not included in the regression.

Table 5 Determinants of the probability to publish

	[1]		[2]		[3]		[4]	
	dF/dx	t value	dF/dx	t value	dF/dx	t value	dF/dx	t value
Gender	0.075**	3.41	0.125**	4.59	0.135**	4.57	0.171**	5.70
Tenured	−0.163**	−6.10	−0.174**	−5.54	−0.177**	−5.27	−0.154**	−4.49
Economics	0.080**	2.78	0.134**	3.74	0.187**	4.83	0.195**	4.92
Econometrics	0.092*	1.81	0.190**	3.10	0.268**	4.12	0.238**	3.31
Public econ.	0.027	0.80	0.044	1.01	0.041	0.84	0.050	0.96
Others	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Fac_Economics	0.036*	1.90	0.073**	3.05	0.061**	2.33	0.069**	2.56
North West	0.042*	1.76	0.036	1.16	0.051	1.53	0.053	1.54
North East	0.055**	2.13	0.028	0.88	0.045	1.27	0.007	0.21
Centre	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
South	−0.039	−1.33	−1.000**	−2.66	−0.067*	−1.65	−0.112**	−2.73
Islands	−0.170**	−4.05	−0.202**	−4.01	−0.194**	−3.61	−0.246**	−4.54
NTLecturer	−0.345**	−7.14	−0.484**	−10.22	−0.539**	−11.84	−0.494**	−11.29
N. Obs.	1,620		1,620		1,620		1,620	
Adj. R^2	0.070		0.103		0.121		0.123	
Obs. P	0.812		0.696		0.601		0.522	
Pred. P	0.834		0.714		0.609		0.515	

Dependent variable: at least one JA in 1969–2006, 0 otherwise

Notes: regressions also include a constant term not included here. Standard errors are robust to heteroscedasticity

* Significant at 10 %; ** Significant at 5 %

[1] at least one JA in 1969–2006, 0 otherwise; [2] at least two JA in 1969–2006, 0 otherwise; [3] at least three JA in 1969–2006, 0 otherwise; [4] at least four JA in 1969–2006, 0 otherwise

otherwise has been introduced in order to take into account the "beginning of career" effect.

Table 5 reports the main findings of the analysis. The results suggest that the probability that an economist will publish an Econlit JA is positively influenced by gender, geographic location of the university, and scientific sub-sector. In particular, these findings shows that men tend to be more productive than women and economists working in Northern universities have a higher probability to publish than those located in Sothern universities. This result can be explained on the basis of the traditional socio-economic dualism characterising the Italian economic system. Dualism which is the result of differences in the social capital endowment and in the cultural and historical traditions. Finally, tenure has a negative impact on the probability to publish, while the dummy indicating the position of "untenured lecturer" is negative and statistically significant. This is because the youngest member of the Italian academic community may be will awaiting publication for already submitted papers. In this first econometric exercise, this variable is the only one that allows us to control for the "age" of an individual economist.

In order to check the robustness of these results, we replicate our analysis by taking into account different values of the cut-off. In other words, we consider not only the probability of publishing at least *one* Econlit JA, but also that of publishing at least *two*, *three* or *four* Econlit JA. The main findings of this analysis are reported in columns [2.], [3.] and [4.] of

Table 6 Determinants of scientific productivity: complete sample

	dF/dx	t values
Propensity to co-authorship	2.912**	3.40
Foreign	0.383**	2.25
Gender	0.327**	3.79
Tenured	0.126	1.30
North West	−0.065	−0.56
North East	−0.295**	−2.52
Centre	Ref.	Ref.
South	0.005	0.05
Islands	−0.086	−0.53
Age97_06	−0.052	−0.26
Age87_96	−0.055	−0.46
Age77_86	Ref.	Ref.
Age69_76	0.134	0.84
Inverse Mill's Ratio	−1.359**	−3.61
N. Obs.	1,015	
Hausman (p value)	0.000	

Dependent variable: scientific productivity

Note: regressions also include a constant term not reported here. Standard errors are bootstrapped (50 replications)

* Significant at 10 %;
** Significant at 5 %

Table 5. The sign and the statistical significance of the main explanatory variables do not tend to change. What it emerges is that only the marginal effects of some variables change. For example, our evidence shows that the probability to publish more than one journal article increases for econometricians. For this group of academics, the marginal effects go from 0.092 for just one Econlit JA to 0.238 for four Econlit JA. Finally, it is interesting to note that the marginal effect of tenure does not change.

The determinants of the scientific productivity

In this section firstly we focus on the roles of the relational and then on the positional variables in the scientific productivity of Italian economists. In both cases, an economist's scientific productivity[34]—the dependent variable in these models—is measured as the total number of Econlit *JA* divided by his/her "seniority". Since we do not have information on the biogical age of the economists in our sample, we use the proxy variable of "seniority", computed as the difference between the year of his/her first Econlit publication and 2006.[35]

In order to fully exploit the use of SNA indexes, the econometric analysis is conducted on both the full sample and the MC of the co-authorship network (i.e. the largest group of economists which are connected by collaboration ties), which accounts for 70 % of the whole sample.

---

[34] We should remark that we do not compute the individual share of each co-author in an article, in other words we do not capture the individual contributions' of each economist to a given paper in order not to introduce arbitrary corrections in our regressions given that within certain research fields—such as economic and econometric theory, history of economic thought, economic history—the number of co-authors tends to be systematically lower than in applied economics.

[35] E.g. for an individual *i* who published his/her first article in an Econlit indexed journal in 1985, the "seniority" is calculated as 2006−1985 = 21.

Role of the relational variables in scientific productivity: full sample

In terms of *relational* driving forces, Italian economist's individual productivity is explained by two types of variables: (1) propensity to co-author and (2) propensity to cooperate with "foreign" researchers. The propensity to co-author (*prop_coauth*) is measured as the proportion of co-authored articles on the total number of articles in the Econlit database. This variable ranges between 0 and 1, where 0 is no collaboration and 1 indicates that all papers are co-authored.

Foreign measures the proportion of collaboration with foreign economists, i.e. set **F**, and is a proxy for the economist's level of "external" connections. For this reason, *Foreign* is calculated as the proportion of foreign economists in total co-authors.[36]

We control for *attributional* variables shown to be significant in the previous econometric exercise (i.e. gender, tenure, geographical location, see Table 5). We control also for age class. Since the 1970s, the university system in Italy has experienced major changes to the career advancement system, research assessment, propensity for internationalization, etc. we try to capture these institutional changes through four time dummies to indicate the date of the first Econlit JA (1997–2006; 1987–1996; 1977–1986 and 1969–1976): *Age97_06, Age87_96, Age77_86,* and *Age69_76*. These dummies control for the different "publication regimes" applying to the economists in our sample.

The econometric analysis involves two econometric problems: (1) sample selection and (2) endogeneity. To address these simultaneously, we estimate a two-stage model. In the first stage, we eliminate the selection bias problem. Selection bias may occur because the average characteristics of the economists in the sample of publishing authors differ from those of the whole population. Without addressing this selection effect, the statistical association between scientific productivity and co-authorship will be inferred incorrectly because the impact of cooperation might be confounded with the coefficients determining the selection. To model the selection mechanism to enable adjustment of the parameter estimations in the structural equation, we adopt a Heckman (1979) procedure. This method uses all the available observations to estimate a *probit* model of the selection indicator. The residuals of this regression are then used to construct a selection bias control factor λ_i: the inverse Mill's ratio which, for each individual, can be computed as:

$$\lambda_i = \frac{\varphi_i(X)}{\Phi_i(X)} \tag{11}$$

where X denotes the covariates used in the selection process, φ_i is the density probability function, and Φ_i is the cumulative probability function. This factor, which accounts for the effects of all the unmeasured characteristics related to the selection variable, is then introduced into the structural equation as an extra explanatory variable. Operationally, we estimate the inverse Mill's ratio from the fourth specification reported in Table 6.

At the second stage, we estimate a *structural equation* given by:

$$prod_i = \alpha_o + \alpha_1 prop\_coauth_i + \alpha_2 foreign_i + X_i'\beta + \theta\lambda_i + \upsilon_i \tag{12}$$

where $prod_i$ is the scientific productivity of each economist i, $prop\_coauth_i$ and $foreign_i$ denote the respective propensity to co-author and to cooperate with foreign researchers; λ_i is the inverse Mill's ratio computed from the previous selection equation and is used to control for selection bias; X indicates other covariates such as *Gender, Tenured,* and

---

[36] Calculated to distinguish each person's identity.

controls such as *North West*, *North East*, *Centre*, *South*, *Islands*, *Age97_06*, *Age87_96*, *Age77_86*, *Age69_76*. Finally, υ_i is the error term with the usual statistical properties.

The direction of causality is an important issue in analysing the relationship between scientific productivity and co-authorship. While it has been shown that cooperation affects productivity, it may well be the case, as suggested by anecdotal evidence, that productivity might affect cooperation: i.e., very productive economists are seen as potentially better co-authors, thus generating a classic reverse causality problem. Also, the presence of endogeneity—i.e., one or more explanatory variables correlated to the true (but unobserved) error term—can generate biased and inconsistent OLS estimates of the coefficient under investigation. We deal with this problem by adopting an IV strategy. The variable that satisfies these conditions is the number of CVA—i.e. chapters in edited books—authored by each economist. CVA are used as an instrument for the propensity for co-authorship because the most recent literature and university policy recommendations assume CVA are the effect of personal connections and relational attitudes that, while not comparable in quality with JA, may reflect alternative use of a researcher's time. If a researcher is contributing to a collected volume, this leaves less time to write JAs. Thus CVA may measure the propensity to cooperate and interact with the wider scientific community irrespective of the impact in terms of the most known bibliometric indexes of scientific productivity.

Under the null hypothesis that the model is appropriately specified with all explanatory variables exogenous, both Hausman tests reject the null hypothesis of exogeneity of the co-authorship propensity variable, which suggests that IV method is the most appropriate method to estimate our model (see the *p* value of the Hausman test in Table 6).

Note that the inverse Mill's ratio variable is always negative[37] and often statistically significant. This suggests that correcting for sample selection is the right choice.

Table 6 reports the main results. The findings suggest that co-authorship is a significant determinant of scientific productivity. In fact, the coefficient of the variable *prop_coath* is positive and statistically significant. This means that economists that are more collaborative are also more productive in JA terms. The results underline another positive role of networking outside the national borders. Collaboration with foreign economists has a positive impact on scientific productivity. Belonging to an international network can thus be interpreted as "signalling" the intrinsic "quality" of the economist and his/her positive political and social attitudes to forging scientific relationships with foreign groups.[38] *Tenure*, *Age* and geographical dummies are in general, not significantly different from zero, with the exception of *North East*.

The role of positional variables in scientific productivity: main component only

Finally, we test for the role of some positional variables characterizing the MC, in explaining the scientific productivity of Italian economists, defined as before. The idea behind the choice of this "biased" sub-sample is to analyse the complex relationships existing between collaborative behaviours and productivity by focussing on the most connected community of economists (i.e. those who are in the "thick of things") so to

---

[37] When the coefficient on the inverse Mill's ratio is negative this means that there are unobserved variables increasing the probability of selection and the probability of a lower than average score on the dependent variable.

[38] Even if we cannot exclude that the liaison with members of the editorial boards of international journals might also play an important role in the publication processes.

avoid the misleading influences of "lone wolves" (people writing alone) and of isolated and marginal couplets or triplets (peripheral cliques).

The variables added to the econometric analysis, derived from the SNA exercise presented in section "Relational structure of co-authorships",[39] are:

1. the betweenness centrality index (*Betweennees*), which measures the strategic influence of an author with special reference to his "bridging role" in relation to different academic groups;
2. the closeness centrality index (*Closeness*), which measures the global centrality of a researcher, i.e. his/her capacity to reach all other researcher through the lowest number of co-authors;
3. the clustering coefficient[40] (CC), which measures the proportion of researcher's co-authors who also co-author with one another;[41]
4. a measure for the stability of the scientific cooperation (*Stability*) which ranges between 0, if all co-authors are different, and 1, if all co-authors are the same.

We also introduce the set of covariates defined previously.[42]

We adopt the econometric approach described in Eq. 12 with the only difference that in this case the estimator is OLS. This choice is justified by the fact that here endogeneity is less a problem because all network indexes refer to second order network features, which are not easily observable by the individual author when deciding on his/her co-authorship strategy.[43] The main findings of our analysis are presented in Table 7.

The variables *Foreign* and *Stability* have sign, size and significance similar to those recorded in Table 5, thus signaling the robustness of our result to the use of different estimation methods.

Both variables referring to the individual economist's centrality (*Closeness* and *Betweenness*) are positive and statistically significant. Centrality, however defined, in the co-authorship network boosts scientific productivity. However, if scientific productivity is the researcher's main goal, then acting as a "bridge" between different scientific sub-sectors and/or different "schools" and academic groups (as measured by betweenness centrality index) is more beneficial than being globally central (as measured by closeness centrality index). Our results are thus in line with a large stream of literature (Burt 1992 and 2004; Fleming et al. 2007) suggesting that network brokers, that span structural holes, enjoy larger than average benefits with respect to other network members since they act as an intermediary between otherwise unconnected sub-groups.

---

[39] These values are calculated exclusively on the MC of the co-authorships network dichotomized according to a threshold value greater than zero.

[40] This value is calculated for each individual based on the aggregated network of the Italian economist community for the whole period 1969–2006.

[41] In general SNA terms, if the neighbors of a single node are also neighbors of each other (i.e. if the clustering coefficient of that node is high), these neighbors are not reliant on the single node for their connection. Therefore, the intermediate node is completely needless for the connection between these two neighbors.

[42] These are *Gender, Tenured*, and controls such as *North West, North East, Centre, South, Islands, Age97_06, Age87_96, Age77_86, Age69_76*.

[43] This is one of the reasons for omitting the most common measure of centrality (degree centrality index) in our case equal to the number of previous co-authors, which is easily observed by the individual researcher: the other reasons are multicollinearity with other measures of centrality.

Table 7 Determinants of scientific productivity (MC)

	[1]		[2]		[3]		[4]	
	Coeff.	*t* value	Coeff.	*t* value	Coeff.	*t* value	Coeff.	*t* value
CC	−0.183**	−3.15	−0.248**	−4.39	−0.095	−1.58	−0.130**	−2.12
Foreign	0.440**	3.46	0.449**	3.50	0.355**	3.15	0.368**	3.82
Stability	0.290**	8.85	0.303**	8.38	0.307**	7.85	0.310**	9.76
Closeness	…	…	2.881**	7.02	…	…	1.055**	2.90
Betweenness	…	…	…	…	4.198**	13.55	3.680**	9.17
Gender	0.277**	4.90	0.241**	5.00	0.242**	5.58	0.234**	4.84
Tenured	0.010	0.18	−0.017	−0.29	−0.044	−0.83	−0.047	−0.87
Geogr. dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age dummnies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mill's Ratio	−0.453**	−2.09	−0.224	−1.00	−0.250	−1.05		
N. Obs.	661		661		661		661	
Adj. R^2	0.248		0.320		0.403		0.409	

Dependent variable: scientific productivity

Notes: regressions also include a constant term not included here. Standard errors are bootstrapped (50 replications)

* Significant at 10 %; ** Significant at 5 %

Clustering is negative and, in most cases, statistically significant, thus signalling the interests of "star" authors in hindering the interactions among their co-authors in order to preserve a hierarchical and productive "hub and spoke" co-authorship structure. This result indirectly confirms the results obtained by Goyal et al. (2006) when looking at the world population of economists,[44] thus showing that, at least in this aspect, the Italian community of economists does not behave differently from other nationality.

Although there are plenty of theoretical reasons for and against frequent changes of co-authors (based on the trust-building vs intellectual novelty trade-off), our econometric results shows that, at least for Italian economists, *Stability* pays. Keeping the same group of authors for successive JA is the best strategy to improve scientific productivity, suggesting that the reduction in transaction costs and the incremental trust building, achieved by repeated interaction prevails over the diversification advantages derived from writing papers with new co-authors.

Conclusions

This paper analyses the effects of the collaborative behaviour (co-authorship) of Italian academic economists on their scientific productivity. To achieve this aim we built an original database by merging the CINECA-MIUR personnel dataset with the Econlit bibliographic dataset. SNA techniques allowed us to study the structure and evolution of

---

[44] "Thus the most connected individuals collaborated extensively and most of their co-authors did not collaborate with each other. The most connected individuals can be viewed as 'stars' from the perspective of the network architecture" (Goyal et al. 2006, p. 412).

co-authorship among Italian economists and to derive positional and relational data for each scientist, exploited in several econometric exercises to explain the scientific productivity (measured by number of journal articles published per year) of this community.

Collaborative behaviour has become more and more frequent within this scientific community as time passed. The resulting network of collaboration is complex (composed of many subnets) and characterized by low density. It is interesting to note that the percentage of foreign economists is higher in the MC than in the whole network, which may signal a correlation between the propensity to co-authorship and the international openness of the Italian academic economist. There are several criteria that might explain the structure and composition of specific sub-networks (such as scientific discipline, geographical location, schools of thought), but the overall structure of the MC displays a small world structure, and the evolution of the network through time is guided by a (non linear) preferential attachment mechanism.

The econometric analysis shows that 'attributional' (age, gender, academic position, tenure, scientific sub-discipline, geographical location), 'relational' (propensity to cooperate and stability of cooperation patterns) and 'positional' (betweenness and closeness centrality indexes and clustering coefficient) variables influence the individual researcher's scientific productivity. It is interesting to note that the econometric results show that the individual productivity of an Italian economist depends (among other factors) on his/her propensity to collaborate, his/her 'international' connections and the stability of his/her collaborative behaviour.

Further, we found that the position of an individual economist in the co-authorship networks affects his/her scientific productivity. Being 'central' increases the number of scientific publications per unit of time, but being a 'bridge' (i.e. connecting two almost separated parts of a network) is even more beneficial.

Finally we showed that stable partnership with other fellows economists, belonging to different cliques (these being geographically, ideologically or subject based), spurs the individual productivity as well as a crafted strategy of co-authorship.

In conclusion our analysis shed some lights on the determinants of the individual scientist productivity as being the results of a number of different individual, institutional and social characteristics.

Future extensions of the paper will use a "quality adjusted" output measures (i.e. numbers of published papers weighted by their impact factors); extend the full sample to "foreign affiliated" Italian academics which, in this case, have been considered only as co-authors; focus on the determinants of co-authorships of a given couplet of researchers.

This paper was originated by pure scientific curiosity and the desire to contribute to the public debate on the university reform process in Italy. Our results suggest that policy makers should be very careful in designing and implementing an explicit incentives system (via a research assessment exercise mainly based on quantitative bibliometric indexes) since the performance of an individual scientist may be influenced by a number of exogenous factors which are beyond his/her control.

Appendix

See Tables 8 and 9.

Table 8 Academic position, scientific sector and gender of Italian Economist as at December the 31st 2006

Scientific Secor (SECSP)	Gender	P	L	SL	AP	CAP	STP	FU	Total
Political economy (01)	Female	1	45	52	22	41	13	27	201
	Male	10	79	53	69	102	54	242	609
Total 01		11	124	105	91	143	67	269	810
Economic policy (02)	Female	1	10	30	10	24	10	13	98
	Male	2	27	24	21	46	15	110	245
Total 02		3	37	54	31	70	25	123	343
Public economics (03)	Female	…	11	15	5	6	4	10	51
	Male	3	16	14	4	22	12	78	149
Total 03		3	27	29	9	28	16	88	200
History of economic thought (04)	Female	…	…	5	…	1	3	2	11
	Male	1	3	4	3	10	2	13	36
Total 04		1	3	9	3	11	5	15	47
Econometrics (05)	Female	…	3	3	3	2	1		12
	Male	…	5	5	3	6	7	23	49
Total 05		…	8	8	6	8	8	23	61
Applied economics (06)	Female	..	5	10	5	8	1	7	36
	Male	1	10	7	22	28	15	40	123
Total 06		1	15	17	27	36	16	47	159
Total economics	Female	2	74	115	45	82	32	59	409
	Male	17	140	107	122	214	105	506	1211
Total		19	214	222	167	296	137	565	1,620

Table 9 Descriptive statistics of dependent variables and regressors

	N. Obs.	Mean	Std. Dev.	Min	Max
Pro_ja	1,317	0.488	0.420	0.027	4.571
Pub_ja	1,620	0.812	0.390	0	1
Prop_coauth	1,317	0.472	0.349	0	1
Non-MIUR econ.	1,015	0.170	0.289	0	1
Stability	1,317	0.448	0.394	0	1
Clustering	1,317	0.144	0.271	0	1
Betweenness	1,034	0.066	0.217	0	1
Closeness	1,034	0.333	0.357	0	1
Gender	1,620	0.747	0.434	0	1
Tenured	1,620	0.680	0.466	0	1
North-West	1,620	0.251	0.434	0	1
North-East	1,620	0.208	0.406	0	1
South	1,620	0.230	0.421	0	1
Fac_economics	1,620	0.611	0.487	0	1
Economics	1,620	0.711	0.453	0	1
Public econ.	1,620	0.123	0.329	0	1
Econometrics	1,620	0.037	0.190	0	1

References

Acedo, F. J., Barroso, C., Casanueva, C., & Galàn, J. L. (2006). Co-authorship in management and organizational studies: An empirical and network analysis. *Journal of Management Studies, 43*(5), 957–983.

Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*(1), 47–99.

Baccini, A., & Barabesi, L. (2010). Interlocking editorship. A network analysis of the links between economic journals. *Scientometrics, 82*(2), 365–389.

Barnett, A. H., Ault, R. W., & Kaserman, D. L. (1988). The rising incidence of co-authorship in economics: Further evidence. *Review of Economics and Statistics, 70*(3), 539–543.

Borgatti, S. P., & Everett, M. G. (1997). Network analysis of 2-mode data. *Social Networks, 19*(3), 243–269.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, Mass: Harvard University Press.

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology, 110*(2), 349–399.

Butler, R. J. (2007). JRI, JF, and the Internet: Coauthors, new authors, and empirical research. *Journal of Risk and Insurance, 74*(3), 713–737.

Cainelli, G., de Felice, A., Lamonarca, M., & Zoboli, R. (2006). The publications of Italian economists in ECONLIT. Quantitative assessment and implications for research evaluation. *Economia politica. Journal of Analytical and Institutional Economics, 23*(3), 385–423.

Cainelli, G., Maggioni, M. A., Uberti, T. E., & de Felice, A. (2012). Coauthorship and productivity among Italian economists. *Applied Economic Letters, 19*(16), 1609–1613.

Dolado, J. J., Garcia-Romero, A., & Zamarro, G. (2003). Publishing performance in economics: Spanish rankings (1990–1999). *Spanish Economic Review, 5*(4), 85–100.

Doreian, P., Batagelj, V., & Ferligoj, A. (2004). Generalized blockmodeling of two-mode network data. *Social Networks, 26*(1), 29–53.

Durden, G. C., & Perri, T. J. (1995). Coauthorship and publication efficiency. *Atlantic Economic Journal, 23*(1), 69–76.

Erdős, P., & Rényi, A. (1959). On random graphs. *Publicationes Mathematicae, 6*, 290–297.

Fafchamps, M., Goyal, S., & van der Leij, M. J. (2010). Matching and network effects. *Journal of the European Economic Association, 8*(1), 203–231.

Fafchamps, M., van der Leij, M.J., & Goyal, S. (2006). Scientific networks and co-authorship. Economics. Oxford: Working Papers, University of Oxford, p. 256.

Fisher, B. S., Cobane, C. T., Ven Vander, T. M., & Cullen, T. M. (1998). How many authors does it take to publish an article? Trends and patterns in political science. *Political Science and Politics, 31*(4), 847–856.

Fleming, L., Mingo, S., & Chen, D. (2007). Brokerage and collaborative creativity. *Administrative Science Quarterly, 52*(3), 443–475.

Freeman, L. C. (1979). Centrality in social networks. Conceptual Clarification. *Social Networks, 1*(3), 215–239.

Goyal, S. (2005). Strong and weak links. *Journal of the European Economic Association, 3*(2–3), 608–616.

Goyal, S. J., van der Leij, M., & Moraga-Gonzalez, J. L. (2006). Economics: An emerging small world. *Journal of Political Economy, 114*(2), 403–432.

Heckman, J. J. (1979). Sample selection as a specification error. *Econometrica, 47*(1), 153–161.

Hilmer, C. E., & Hilmer, M. J. (2005). How do journal quality, co-authorship, and author order affect agricultural economists' salaries? *American Journal of Agricultural Economics, 87*(2), 509–523.

Holder, M., Langrehr, F., & Schroeder, D. (2000). Finance journal co-authorship: How do co-authors in very selected journals evaluate the experience? *Financial Practice and Education, 10*(1), 142–152.

Hollis, A. (2001). Co-authorship and the output of academic economists. *Labour Economics, 8*(4), 503–530.

Hudson, J. (1996). Trends in multi-authored papers in economics. *Journal of Economic Perspectives, 10*(3), 153–158.

Johnston, D. W., Piatti, M., & Torgler, B. (2013). Citation success over time: Theory or empirics? *Scientometrics, 95*, 1023–1029.

Kalaitzidakis, P., Mamuneas, T. P., & Stengos, T. (1999). European economics: An analysis based on publications in the core journals. *European Economic Review, 43*(4–6), 1150–1168.

Kretschmer, H. (1994). Coauthorship networks of invisible colleges and institutionalized communities. *Scientometrics, 30*(1), 363–369.

Laband, D. N. (1987). A Quantitative test of journal discrimination against women. *Eastern Economic Journal, 13*, 149–154.

Laband, D., & Piette, M. J. (1995). Team production in economics: Division of labour or mentoring. *Labour Economics, 2*(1), 33–40.

Laband, D., & Tollison, R. (2000). Intellectual collaboration. *Journal of Political Economy, 108*(3), 632–661.

Laband, D., & Tollison, R. (2006). Alphabetized co-authorship. *Applied Economics, 38*(14), 1649–1653.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673–702.

Liebowitz, S. J., & Palmer J. P. (1983). Assessing assessments of the quality of economics departments, manuscript. Rochester, N.Y.: University of Rochester.

Maggioni, M. A. (1995). Metodologie reticolari per l'analisi dei sistemi locali di produzione e innovazione. *Economia and Lavoro, 29*(1/2), 109–130.

Maggioni, M. A., Gambarotto, F., & Uberti, T. E. (2009) Mapping the Evolution of Clusters. A Meta-Analysis. Milano: Working Papers, Fondazione Eni Enrico Mattei, p. 74.

Maggioni, M. A., & Uberti, T. E. (2009). Knowledge networks across Europe: Which distance matters? *Annals of Regional Science, 43*(3), 691–720.

Maggioni, M. A., & Uberti, T. E. (2011). Networks and geography in the economics of knowledge flows. *Quality & Quantity, 45*(5), 1031–1051.

McDowell, J. M., & Melvin, M. (1983). The determinants of co-authorship: An analysis of the economics literature. *Review of Economics and Statistics, 65*(1), 155–160.

McDowell, J. M., & Smith, J. (1992). The effect of gender-sorting on propensity to coauthor: Implications for academic promotion. *Economic Inquiry, 30*(1), 68–82.

Medoff, M. H. (2003). Collaboration and the quality of economics research. *Labour Economics, 10*(5), 597–608.

Neary, J. P., Mireless, J. A., & Tirole, J. (2003a). Evaluating economics research in Europe: An introduction. *Journal of the European Economic Association, 1*(6), 1239–1249.

Neary, J. P., Mireless, J. A., & Tirole, J. (2003b). Evaluating economics research in Europe: An introduction. *Journal of the European Economic Association, 1*(6), 1239–1249.

Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences, 98*(2), 404–409.

Newman, M. E. J. (2003). The structure and function of complex networks. *Society for Industrial and Applied Mathematics Review, 45*(2), 167–256.

Nowell, C., & Grijalva, T. (2011). Trends in co-authorship in economics since 1985. *Applied Economics, 43*(iss. (28-30)), 4369–4375.

Piette, M. J., & Ross, K. L. (1992). An analysis of the determinants of co-authorship in economics. *Journal of Economic Education, 23*(3), 277–283.

Sauer, R. (1988). Estimates of the returns to quality and co-authorship in economic academia. *Journal of Political Economy, 96*(4), 855–866.

Sigelman, L. (2009). Are two (or three or four … or nine) heads better than one? Collaboration, multi-disciplinarity, and publishability. *Political Science and Politics, 42*(3), 507–512.

Strogatz, S. (2001). Exploring complex networks. *Nature, 410*, 268–276.

Sutter, M., & Kocher, M. (2004). Patterns of co-authorship among economists in the USA. *Applied Economics, 36*(4), 327–333.

Ter Wal, A. L. J., & Boschma, R. A. (2009). Applying social network analysis in economic geography: Framing some key analytic issues. *Annals of Regional Science, 43*(3), 739–756.

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology, 111*(2), 447–504.

van der Leij, M., & Goyal, S. (2011). Strong ties in a small world. *Review of Network Economics, De Gruyter*, 10(2), 1–22.

Vieira, P. C. C. (2008). An economics journals' ranking that takes into account the number of pages and co-authors. *Applied Economics, 40*(7), 853–861.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

Watts, D. (1999). *Small worlds: The dynamics of networks between order and randomness*. Princeton: Princeton University Press.