Economic Modelling xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Economic Modelling

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

Co-authorship and research productivity in economics: Assessing the assortative matching hypothesis

Damien Besancenot[a,1], Kim Huynh[b,2], Francisco Serranito[c,*,3]

[a] Université Paris Descartes, Centre Universitaire des Saints Pères, 45 rue des Saints Pères, 75270 Paris cedex 06, France
[b] Université Panthéon-Assas Paris 2, 4 Rue Blaise Desgoffe, 75006 Paris, France
[c] Université Paris 13, Sorbonne Paris Cité, 99, av. J.-B. Clément, 93430 Villetaneuse, France

ARTICLE INFO

JEL codes:
A14
C25
D83
I23

Keywords:
Co-authorship
Count data
Zero inflate models
Instrumental variables
h index

ABSTRACT

This paper estimates the relation between the size and quality of scientists' co-author networks and individual characteristics (notably productivity) in the context of institutional changes in French academia in the mid-1980s. The analysis employs the Two-Stage Residual Inclusion (2SRI) framework to handle endogeneity in individual productivity relative to the quality of co-authors. The main finding is that the size and quality of authors' networks are positively related to their productivity; this is understood as evidence of assortative matching. Other effects on co-author networks (such as life-cycles, specialties fields) are also identified. Our results have important policy implications as it indirectly demonstrates the effectiveness of career incentives linked to publication.

1. Introduction

In its traditional approach, sociology of science considers research as a solitary activity. Science is a winner takes all race and all the prestige of a breakthrough is granted to the first author who publishes a new result. Co-authorship, which limits the prestige of being at the origin of a new concept, seems thus unnatural (Stephan, 1996). Moreover, co-authorship requires coordination efforts, imposes compromises between authors or limits the innovative content of the collective work when authors present different degree of risk aversion (Hudson, 1996). Co-authorship may also suffer from free riding, the workload may be unfairly distributed and it is difficult ex ante to identify the true ability of a potential coauthor (Hollis, 2001; Fafchamps et al., 2010). Last but not least, the fact that an author writes only co-authored papers is usually seen by his peers as a signal of his inability to write a paper alone and is therefore considered as a negative signal.

For all these reasons, it seems difficult to understand the monotonic long run increase in co-authorship recorded by top tier journals. In their study of *The American Economic Review*, Torgler and Piatti (2013) remark that while 99% of the papers were single authored in the period between 1911 and 1920; the number falls to only 28% during the 2001 to 2010 period. Moreover, there is a constant increase in the number of coauthors; the frequency of papers authored by three or more authors rose from less than 1% during the 1960s to 22% by the 2001 to 2010 period. Hamermesh (2013) also notes that after the apparition in 1993 of the first four-authored papers in the American Economic Review, the Journal of Political Economy and the Quarterly Journal of Economics, the first five and six-authored papers start to appear since the beginning of the 2010s. The same evolution is put forward by Card and DellaVigna (2013) who stress that among the papers published in the 'top five' economic journals, the average number of authors rose from 1.3 in the early 1970s to 2.2 in 2011–2012.

Beginning at the end of the 1970s, a vast strand of research documents that co-authorship is no longer the exception but constitutes a new scientific norm (Beaver and Rosen, 1978; Stefaniak, 1982; Zitt et al., 2000; Laband and Tollison, 2000; Hamermesh, 2013, 2015). Generally, the economic literature explains this evolution through the positive effects of scientific collaboration on the quantity and the quality of the research output. With respect to quantity, co-authorship is a simple way to increase the number of papers that a researcher is

* Corresponding author.
E-mail addresses: damien.besancenot@parisdescartes.fr (D. Besancenot), kim.huynh@u-paris2.fr (K. Huynh), serranito@univ-paris13.fr (F. Serranito).
[1] LIRAES *EA 4470*.
[2] LEMMA EA 4442.
[3] CEPN (UMR CNRS 7234) and IRD, LEDa, DIAL UMR 225.

http://dx.doi.org/10.1016/j.econmod.2017.05.018
Received 27 December 2016; Received in revised form 19 April 2017; Accepted 30 May 2017

able to publish in a given time period. As two two-authored papers are generally worth more than one single-authored paper, the incentives to collaborate are strong (Barnett et al., 1988). Therefore, various empirical works seem to confirm the link between the volume of publication and collaboration. Durden and Perry (1995) and Lee and Bozeman (2005) find that the total number of publications is significantly and positively related to the number of collaborative publications. Hollis (2001) shows that the more co-authorship done in the past, the more prolific an author is likely to be today and Lee and Bozeman (2005) stresses that collaboration is a strong predictor of the total number of a researcher's publications. Empirical evidence concerning the link between research quality and co-authorship is less conclusive, particularly if quality is measured by the number of citations (see Levitt (2015) for a review). On the one hand, Laband (1987), Johnson (1997), and more recently Levitt (2015) report that co-authored papers are cited more often than single-authored ones; but on the other hand, Barnett et al. (1988) and Hollis (2001) find no effect. Chung et al. (2009) show that papers co-authored with a prolific author receive more citations, whereas papers written with colleagues from the same institution does not disclose any increase in quality.

As co-authorship plays a central role in the production of knowledge, it is essential to understand the factors that favor or hinder collaboration. However, if the reason why researchers decide to collaborate has been largely studied (Beaver, 2001 listed 18 motives justifying scientific collaboration[4]), few research has been devoted to understand the specific characteristics which lead academics to choose each other in the building of a new team.

If we exclude the specific cases where collaboration is justified by friendship or is considered as a way to escape academic isolation (Medoff, 2003; Acedo et al., 2006; Hamermesh, 2013), team formation is mainly explained by advocating the role of complementarities in researchers' abilities. In a pioneering paper, McDowell and Melvin (1983) linked the rise in co-authorship to the explosion of knowledge in economics. While researchers are involved in increasing specialization, working with colleagues with complementary skills allows alternative approaches, division of labor and improves the academic production. Under alternative presentations, this seminal argument has been developed in a series of contributions. For instance, authors who work in areas outside of their specialty tend to engage more in co-authorship than authors with close scientific tools (Piette and Ross, 1992). More recently, Chan et al. (2016) focused on Nobel Laureates and showed that scientific collaboration is fostered by conceptual complementarities, which erode over time following repeated interactions.

Strategic behaviors may also influence team formation. For instance, Ong et al. (2015) show that co-authorship may be affected by the order in which authors are listed on the title page. As authors with earlier last names initials have better visibility, they are therefore more keen to start collaborations.

Beside these arguments, Fafchamps et al. (2010) focus on the role of ability in the matching of co-authors. They argue that research collaboration is most likely between authors with similar abilities – *the assortative matching hypothesis*[5]. Initially, assortative matching has been studied in human mating by Pearson (1903) who reported strong correlations in physical appearance (height, span of arms or length of forearms) between husbands and wives. In social relationship, assortative matching occurs over a wide array of human characteristics such as socio-economic status, educational level, religious or political attitudes. In co-authorship, assortative matching may be observed between researchers exhibiting the same characteristics.

Putting forward evidences of gender sorting in team formation, Boschini (2007) underlines that the propensity to co-author with a woman is higher for women than for men. Moreover, this propensity gap increases with the presence of women in the field of research. This result is consistent with experimental evidence that the gender composition of teams affects team productivity (Ivanova-Stenzel and Kubler, 2011). The fact that women seem to perform worse in case of gender mixed teams may explain their under representation in research activities which in turn leads to a lower rate of co-authorship for female academics. Age of the authors is also relevant in co-authorship as revealed by Hamermesh (2015). For young or mid-career researchers, the average coauthor is the scholar's contemporary. A similarity that disappears while considering researchers aged over 51 who seem to prefer co-authoring with mid-career colleagues. Assortative matching with respect to ethnicity is put forward by Freeman and Huang (2014) who show that author with similar ethnicity develop more collaboration than expected given proportion among authors. Finally, in a survey of 580 economists from 69 countries, Kumar and Ratnavelu (2016) note that more than 21.5% of these researchers prefer working with colleague belonging to his/her department most or all of the time.

Our paper contributes to this literature by empirically assessing Fafchamps et al. (2010) matching hypothesis. We estimate the link between the "individual research quality" of an academic and that of his/her co-authors. Determinants of co-authorship are estimated based on a novel database considering all academic economists with a position in a French university in 2004. Our dataset has two main advantages: it is exhaustive and it includes both publishing and non-publishing academics. In general, studies applying bibliometrics never include the second category of academics. Our paper employs a specific econometric framework in order to take account of these non-publishing academics; so it is likely to produce more trustworthy estimates.

As our purpose is to assess the link between the productivity of a researcher and the quality of his network of co-authors, a key feature of our approach is the measure of individual productivity. In the scientometric literature, researcher's productivity is assessed according to the various characteristics of the academic output. One may focus on the number of author's papers, the citations received by each paper, a normalized measure of these citations taking into account the citing journal or any combination of these measures such as the h or the g indexes (Hirsch, 2005; Egghe, 2006). One may also consider the quality of the publication media in which papers have been published. This quality can be measured by indexes such as the Journal Impact Factor (Garfield, 1972), the Audience Factor (Zitt and Small, 2008), the Eigenfactor (Bergstrom, 2007), or the place in journal ranking (e.g. Harzing's Journal Quality List, 2017). Each of these measures reflecting a partial information on an author's performance, a comprehensive assessment of the overall research productivity of an academic requires to integrate these multiple indicators into a single composite index (see for instance Sahoo et al. (2017) who computed a composite research productivity index mixing six different indicators with a weighting system defined through a directional Benefice of Doubt model).[6]

However, in this paper, we will consider separately the elements that reflect the reputation of the academic (the number of published papers and the citations they receive) and the quality measure of the journals in which the academic output has been published. This

[4] Access to expertise, better access to resources and funds, to obtain prestige, learning technical knowledge, progressing more rapidly, tackling more difficult problems, enhancing productivity, creating a network, to break into a new field, satisfying curiosity, sharing the excitement of an area with other people, reducing errors, staying focused on research, reducing isolation, education (student education), advancing knowledge and finally for fun.

[5] Collaboration between authors with different abilities can only happen if the contribution of the lower-ability author relaxes the time-constraints of his/her co-authors. In this case, collaboration enables higher-ability authors to produce more research, while lower-ability researchers produce better-quality output than they would otherwise.

[6] The input-output production approach developed by Lee and Worthington (2016) to assess Australian universities may also be used to measure research performance at the individual level. In this approach the efficiency score is built taking into account the role of academic publications in the various other researcher's production. Such measures allow taking into account a wider dimension of researchers' production and productivity.