



ELSEVIER

Contents lists available at ScienceDirect

Journal of Informetrics

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

A regression analysis of researchers' social network metrics on their citation performance in a college of engineering



Oguz Cimenler^{a,*}, Kingsley A. Reeves^a, John Skvoretz^b

^a Industrial & Management Systems Engineering, University of South Florida, 4202 East Fowler Avenue, ENB118, Tampa, FL 33620, United States

^b Sociology, University of South Florida, 4202 East Fowler Avenue, Tampa CPR 107, FL 33620, United States

ARTICLE INFO

Article history:

Received 10 March 2014

Received in revised form 4 June 2014

Accepted 5 June 2014

Available online 28 June 2014

Keywords:

Collaborative networks

Social network analysis

Poisson regression

Self-reported data

Citation-based research performance

ABSTRACT

Previous research shows that researchers' social network metrics obtained from a collaborative output network (e.g., joint publications or co-authorship network) impact their performance determined by g-index. We use a richer dataset to show that a scholar's performance should be considered with respect to position in multiple networks. Previous research using only the network of researchers' joint publications shows that a researcher's distinct connections to other researchers, a researcher's number of repeated collaborative outputs, and a researchers' redundant connections to a group of researchers who are themselves well-connected has a positive impact on the researchers' performance, while a researcher's tendency to connect with other researchers who are themselves well-connected (i.e., eigenvector centrality) had a negative impact on the researchers' performance. Our findings are similar except that we find that eigenvector centrality has a positive impact on the performance of scholars. Moreover, our results demonstrate that a researcher's tendency toward dense local neighborhoods and the researchers' demographic attributes such as gender should also be considered when investigating the impact of the social network metrics on the performance of researchers.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

It is important to determine who are the most influential researchers and invest in those researchers to both maximize the research outputs and to allocate funding effectively (Abbasi, Altmann, & Hossain, 2011; Jiang, 2008). Influential researchers can be determined by using social network metrics such as centrality metrics after mapping their collaborative output networks (e.g., joint publications, grant proposals, and patents) in which a tie between any two authors indicates collaboration on the making of a collaborative output. Hou, Kretschmer, and Zeyuan (2008) found that there was a positive correlation between being an influential researcher, (i.e., having a high degree centrality in the collaborative output network) and output of a researcher (i.e., number of publications). Defazio, Lockett, and Wright (2009) also found that there was high impact of being an influential researcher in the collaborative output network on output of a researcher. However, the quality of research outputs is as important as the quantity of the research outputs.

Hirsch (2005) proposed an index called the h-index in order to attempt to measure both the number of publications a researcher produced (i.e., quantity) and their impact on other publications (i.e., quality). Using the researchers' publications

* Corresponding author. Tel.: +1 813 974 2269.

E-mail addresses: oguzcimenler@gmail.com, ocimenle@mail.usf.edu (O. Cimenler).

Table 1
Advantages of scientific collaboration.

Access to expertise for complex problems, new resources and, funding	Katz and Martin (1997), Melin (2000), Beaver (2001), Hara et al. (2003), Sonnenwald (2007), Bukvova (2010), National Science Board report (2012) and Hale (2012)
Increase in the participants' visibility and recognition	Katz and Martin (1997) and Beaver (2001)
Rapid solutions for more encompassing problems by creating a synergetic effect among participants	Melin and Persson (1996) and Beaver (2001)
Decrease in the risks and possible errors made, thereby increasing accuracy of research and quality of results due to multiple viewpoints	Beaver (2001) and Bukvova (2010)
Growth in advancement of scientific disciplines and cross-fertilization across scientific disciplines	Beaver (2001) and Cummings and Kiesler (2005)
Development of the science and technical human capital, e.g., participants' formal education and training, and their social relations and network ties with other scientists	Bozeman and Corley (2004)
Increase in the scientific productivity of individuals and their career growth	Fox (1983), Katz and Martin (1997), Bozeman and Corley (2004) and Lee and Bozeman (2005)

data in the information schools of five universities, Abbasi et al. (2011) investigated the impact of social network metrics (including different centrality metrics, average tie strength, and efficiency coefficient proposed by Burt (1992)) obtained from a researchers' co-authorship network on their g-index (another form of h-index), and found out that degree centrality, average tie strength, and efficiency coefficient had a positive impact on the researchers' performance, while eigenvector centrality had a negative impact on the researchers' performance. Their study can be extended by considering the network metrics obtained from researchers' multiple networks. Thus, the purpose of our study is to test the findings of Abbasi et al. (2011) with the social network metrics obtained from researchers' multiple collaborative networks defined by joint publications, joint grant proposals, and joint patents as well as their communication network to understand the relationship between these social network metrics and the performance of researchers. Collecting researchers' ties for their informal conversational exchange (or informal communication) and collaborative outputs with other researchers within a college simultaneously makes this testing possible. We use h-index instead of the g-index because the researchers within the same field of study are compared (Bornmann & Daniel, 2009). In sum, this study seeks an answer to the following question: *what is the impact of social network metrics obtained from researchers' communication and collaborative output networks on their performance as measured by citations of their publications?*

2. Literature review and hypotheses

2.1. Researchers' communication and collaborative output networks

A science and technology (S&T) system comprises a wide range of activities such as fundamental science or scholarly activity, and applied research and developmental activities mainly concentrating on creating new products and processes (Moed, Glänzel, & Schmoch, 2004). It has become a driving force over the last 20 years for major economic growth and development and it is, therefore, an inseparable part of several national and regional innovation systems (Freeman & Soete, 2009; Moed et al., 2004). One of the important attributes contributing to the S&T system performance is scientific collaboration (Hara, Solomon, Kim, & Sonnenwald, 2003; Moed et al., 2004). Sonnenwald (2007) defined scientific collaboration as the interaction within a social context among two or more scientists in order to facilitate the completion of tasks with regard to a commonly or mutually shared goal. Thus, participants in the collaboration event integrate valuable knowledge from their respective domains to create new knowledge. Scientific collaboration provides several salient advantages as shown in Table 1. One of the important factors leading to advantages of scientific collaboration is the social dimension of scientific work such as informal conversational exchanges between colleagues (Bozeman & Corley, 2004; Katz & Martin, 1997), co-authorship relations (Glänzel & Schubert, 2004; Katz & Martin, 1997), jointly submitted grant proposals (Katz & Martin, 1997; Rigby, 2009), and co-patent applications (Balconi, Breschi, & Lissoni, 2004; Breschi & Lissoni, 2004, 2009; Meyer & Bhattacharya, 2004).

Co-authorship in scholarly publications is the most tangible and well-documented forms of scientific collaboration, and it is also a good indicator of the S&T system performance. Therefore, it is used widely in scientific collaboration studies (Glänzel & Schubert, 2004; Katz & Martin, 1997; Melin & Persson, 1996; Moed et al., 2004). For example, using social network analysis (SNA), Newman (2001a, 2001b, 2001c) and Barabasi et al. (2002) analyzed the structural properties of scientific collaboration patterns in large scale by depicting the network of researchers when two authors were considered linked if their names appeared in the same scientific journal. They found that co-authorship networks were small world networks in which most nodes (i.e., authors) could be reached from other nodes by a small number of steps. With a similar approach used in co-authorship network studies, some studies also analyzed the structure of co-inventor maps in the case that two patent applicants (i.e., co-authors) were linked if there was a patent application together by these two applicants; thus, a network of co-invention was constructed. However, analyzing co-inventor maps was not used as widely as analyzing co-authorship maps (Breschi & Lissoni, 2004). In addition, for the networks constructed from researchers' jointly submitted grant proposals,

there was not any study in the literature analyzing the properties of these networks, their relations with other concepts, and related implications.

Many scholars argue that co-authorship alone is insufficient as a measure of research collaboration. For example, [Katz and Martin \(1997\)](#) pointed out that many cases of collaboration did not result in co-authored publications; for example when researchers worked closely together, but decided to publish their results separately due to the fact that they came from different fields and desired to produce single-author papers in their own discipline. Their study concluded that measuring co-authorship was a partial indicator of research collaboration. [Melin and Persson \(1996\)](#) also asserted that co-authorship was only a rough indicator of collaboration, even though significant scientific collaboration leads to coauthored publications in most cases. The qualitative study of [Laudel \(2002\)](#) determined different types of collaborations that were classified according to the content of contribution made by collaborators. Then, a collaborator was rewarded with a co-authorship depending on the level of his/her contribution. The assumption that co-authorship and research collaboration are synonymous was criticized by several other scholars for the following reasons: listing co-authors for purely social reasons ([Bozeman & Corley, 2004](#); [Hagstrom, 1975](#); [Katz & Martin, 1997](#)), listing co-authors simply by the virtue of providing material or performing a routine task ([Bozeman & Corley, 2004](#); [Katz & Martin, 1997](#); [Stokes & Hartley, 1989](#)), making the colleagues 'honorary co-authors' ([Bozeman & Corley, 2004](#); [Katz & Martin, 1997](#); [LaFollette, 1992](#)), and listing co-authors who did not even communicate with each other during research collaboration (e.g., many publications in physics and astrophysics include hundreds of authors) ([Pepe, 2011](#)).

[Fox \(1983\)](#) stated that communication and exchange of research findings and results were the most fundamental social process of science, and the principal means of this communication was the publication process. Communication between researchers not only stimulates them to think regarding the unsolved problems in their fields and possible research projects, thereby developing new ideas and solutions, but it also transmits 'know-how' or the procedural knowledge to efficiently solve the problems to other researchers ([Laudel, 2002](#)). Collaborations mostly begin informally and arise from informal communication between researchers, i.e., through close personal contacts and professional networks ([De Solla Price & Beaver, 1966](#); [Hagstrom, 1975](#); [Edge, 1979](#); [Katz & Martin, 1997](#); [Bozeman & Corley, 2004](#); [Tijssen, 2004](#)). [Kraut and Egido \(1988\)](#) found out that researchers in a close physical proximity tended to collaborate more due to the changes in three properties of informal communication: increasing the frequency of communication, increasing the quality of communication, and reducing the cost of communication. [Olson and Olson \(2000\)](#) also reported that face-to-face communication facilitates the flow of situated cognitive and social activities due to some of its key characteristics such as rapid feedback and multiple channels (e.g., voice, facial expression, gesture, body posture). However, the use of information and communication technologies (ICT) such as audio and video conferences, mobile phones, e-mail, social networking sites especially designed to support collaborative environment, and the World Wide Web facilitate informal communication between researchers and help them collaborate with other distant researchers in a timely manner ([Borgman & Furner, 2002](#); [Schleyer et al., 2008](#); [Sonnenwald, 2007](#)). Using both types of communication, face-to-face and ICT, have their own advantages and disadvantages ([Olson & Olson, 2000](#)). In sum, communication is an important source and influential factor for scientific collaboration ([Bukvova, 2010](#); [Glänzel, 2002](#); [Hara et al., 2003](#); [Katz & Martin, 1997](#)) and a fundamental component to sustain collaboration ([Sonnenwald, 2007](#)).

Many scholars make a clear distinction between researchers' communication and collaboration. For example, [Melin and Persson \(1996\)](#) reported that 'collaboration was an intense form of interaction that allowed for effective communication'. [Melin \(2000\)](#) discussed that collaboration could be measured in a number of ways such as exchange of phone calls and e-mails, but a more concrete form to measure the collaboration was through co-authorship information. [Laudel \(2002\)](#) accepted publications as a way of formal communication, and found out that a considerable proportion of collaborations were not rewarded as a co-authorship. [Borgman and Furner \(2002\)](#) discussed that collaboration was one of the communication behaviors exhibited by authors in their various capacities. Similarly, from a network viewpoint, [Newman \(2001c\)](#) reported that there was an assumption that most people who wrote a paper together might not be genuinely acquainted with one another. Consequently, even though there is a clear distinction between researchers' communication and collaboration, considering the researchers' communication and collaborative output networks separate from each other is not fully addressed in the literature. Taking the assumption reported by [Newman \(2001c\)](#), one notable study made by [Pepe \(2011\)](#) compared the structure of researchers' communication network with the structure of their collaborative output network (e.g., co-authorship network) by utilizing techniques used in SNA. The study found out the extent to which the structure of researchers' communication network overlaps the structure of their collaborative output network. That is, the more these network structures overlap the more likely collaborative output relations between researchers can be seen as a surrogate or proxy for communication relations between researchers.

2.2. A solution to collect researchers' multiple collaborative output networks as well as their communication network

Considering the discussion in the literature that relying solely on co-authorship relations is not a sufficient indicator of scientific collaboration, [Bozeman and Corley \(2004\)](#) and [Lee and Bozeman \(2005\)](#) employed participants' self-report of collaboration information, which permitted the participants to indicate which relationships are worthy of being considered as collaborations. Using a questionnaire, they asked participants to make a self-report of the number of people with whom they had engaged in research collaborations within the past 12 months. Even though [Lee and Bozeman \(2005\)](#) and [Vasileiadou \(2009\)](#) highlighted the disadvantages of the self-reported way of collecting data such as accuracy of the collected data,

there are many recent studies using the method of collecting collaboration information via self-report (Duque et al., 2005; Sooryamoorthy & Shrum, 2007; Van Rijnsoever, Hessels, & Vandenberg, 2008; Ynalvez & Shrum, 2011).

Their method can be extended to collecting researchers' communication and collaboration information in a social network context by employing a questionnaire where researchers identify their contacts and provide the amount of communication and collaboration with those contacts via self-reports. For example, while collecting the collaboration information, a participant can be asked to report the names of the researchers with whom he/she has engaged in both communication and research collaborations together with the frequency of that communication and the number of collaborative outputs (both in-progress and completed) with those reported names via a name generator. By reporting both of their in-progress and completed collaborative output ties (e.g., co-authorship ties), they can decide on which ties are important to them and whether or not reported contact is actually involved in research. This helps overcome the challenge that many collaborations do not result in tangible outcome such as co-authorship by capturing in-progress collaborative output ties as well as other challenges, such as co-authors who are listed for only social reasons and co-authors that are not even communicated. It will be more successful if this method can be executed within the college of a university or even within a university because close proximity of the researchers will facilitate data collection in a way that the relational data for mapping the researchers' multiple networks (e.g., network of communications, network of joint publications, grant proposals, and patents) can be simultaneously collected at either the individual college level or at the university as a whole. Moreover, the name generator can contain prepopulated names of the researchers within the college of a university in order to help the participant for ease of remembering the names.

In addition to abovementioned advantages, administering a self-reported questionnaire can overcome the major limitation in gathering data with regard to a researcher's communication as well as collaborative output information with other researchers. The limitation is mainly due to these challenges: the unavailability of data for multiple networks, the inability to access the multiple data repositories, and the difficulty of scanning multiple databases. For example, for the same researchers, data might be available and easily accessible in order to construct the network of co-authorships or joint publications, but either unavailable or difficult to access in order to construct the network of communications, joint grant proposals, and patents. Moreover, scanning the different databases to collect the same researchers' both communication and collaborative output information might also be tedious job.

2.3. A performance measure of researchers: h-index

A researcher's performance is assessed by two factors: the number of publications he/she produced and the impact of those publications in the scientific community (Bornmann & Daniel, 2007, 2009). Hirsch (2005) proposed an index called h-index that combined both of these quantity and impact factors. The h-index drew the attention of many researchers in the scientific community, and many publications on this topic emerged (Costas & Bordons, 2007). Hirsch (2005) defined the h-index as follows: "A scientist has index h if h of his/her N_p papers have at least h citations each and the other ($N_p - h$) papers have fewer than h citations each, where N_p is the number of papers published over n years" (Cronin & Meho, 2006). Even though the h-index was better than straight citation counts (Cronin & Meho, 2006) and had more predictive power to assess the future achievement of researchers (Hirsch, 2007), different modifications of the h-index have been proposed in the literature to overcome its shortcomings (Bornmann, Mutz, & Daniel, 2008; Costas & Bordons, 2007). Some shortcomings are the followings: favoring disciplines which do experimental research study in larger groups such as physics, assigning an equal value to each author in multiple-author papers, not accounting for author sequence and the total number of authors, being inflated via self-citations, not considering books and other alternative forms of publication, not considering the performance changes throughout a researcher's career and lag time between a paper being published and being discovered and cited (McCarty, Jawitz, Hopkins, & Goldman, 2013). In this study, the h-index, the most widely used performance metric for researchers, was used because the researchers within the same field of study are compared (Bornmann & Daniel, 2009).

2.4. Social network metrics

Sonnenwald (2007) defined scientific collaboration as the interaction within a social context among two or more scientists in order to facilitate the completion of tasks with regard to a commonly shared goal. Thus, those collaborations are perpetuated through social networks (Abbasi et al., 2011). SNA is the method used to reveal the structure of collaboration between individuals (Hou et al., 2008; Kretschmer, 2004). Hence, many social network metrics in SNA are used to analyze the structure of collaboration between researchers (Friedkin, 1978; Newman, 2001a, 2001b, 2001c). Using the data gathered by the questionnaire, our goal in this study is to test the impact of following social network metrics extracted from both researchers' communication and collaborative output networks on the researchers' citation-based performance index (h-index).

- *Degree centrality* (i.e., the researchers' distinct connections to many different researchers)
- *Closeness centrality* (i.e., the shortness of a researcher's total distance to all other researchers)
- *Betweenness centrality* (i.e., the number of times the researchers holding the shortest path between two other researchers)
- *Eigenvector centrality* (i.e., the researcher's tendency to connect with other researchers who are themselves well-connected)

- *Average tie strength* (i.e., the researcher's averaged number of repeated collaborative outputs with other researchers)
- *Burt's efficiency coefficient* (i.e., the researchers' redundant connections to a group of researchers who are themselves well-connected)
- *Local clustering coefficient* (i.e., an researcher's tendency toward the dense local neighborhoods)

Degree centrality of a node n_i , denoted by $C_D(n_i)$, is the number of nodes that adjacent to node n_i or the number of unique edges, e_{ij} , that are connected to node n_i (Wasserman & Faust, 1994). Normalized degree centrality, $C'_D(n_i)$, is found by dividing the degree centrality of node n_i by the number of total nodes, n , excluding n_i such as $(n - 1)$. Then, the normalized degree centrality can be used to compare the degree centrality of nodes across networks of different size. Thus, $C'_D(n_i)$ which ranges from 0 to 1 is given by:

$$C'_D(n_i) = \frac{C_D(n_i)}{n - 1} = \frac{\sum_j e_{ij}}{n - 1}, \tag{1}$$

where $\sum_j e_{ij} = \sum_i e_{ji}$ for undirected networks.

Closeness centrality of a node n_i , denoted by $C_C(n_i)$, is the sum of geodesic distances to all other nodes in a network (Wasserman & Faust, 1994). Geodesic distance is a shortest path (i.e., lowest total number of edges) linking node, n_i and n_j , which is denoted by $d(n_i, n_j)$. Then, the sum of geodesic distances is shown by $\sum_j^n d(n_i, n_j)$. A lower closeness centrality score indicates a more central position for a node in a network (Hansen, Shneiderman, & Smith, 2011). Sabidussi's (1966) index of actor closeness offers the sum of reciprocal geodesic distances (Wasserman & Faust, 1994). Thus, the higher values indicate more central position. The normalized closeness centrality, $C'_C(n_i)$ which ranges from 0 to 1, is found by multiplying $C_C(n_i)$ by $n - 1$. Then, $C'_C(n_i)$ is given by:

$$C'_C(n_i) = \frac{n - 1}{\sum_j^n d(n_i, n_j)}. \tag{2}$$

Betweenness centrality of a node n_i , denoted by $C_B(n_i)$, is the sum of the ratio of the number of geodesics, $g_{jk}(n_i)$, linking the nodes n_j and n_k that contain node n_i to the number of geodesics, g_{jk} , linking the nodes n_j and n_k (Wasserman & Faust, 1994). In other words, it counts “the number of geodesic paths (i.e., shortest paths) that pass through a node n_i (Borgatti, 2005). The more a node is high in betweenness centrality the more a node is in the position to broker information and ideas (McCarty et al., 2013). The normalized betweenness centrality, $C'_B(n_i)$ which ranges from 0 to 1, is found by dividing the betweenness centrality by $(n - 1)(n - 2)/2$, which indicates the number of pairs of nodes not including n_i . Then, $C'_B(n_i)$ is given by:

$$C'_B(n_i) = \frac{C_B(n_i)}{(n - 1)(n - 2)/2} = \frac{\sum_{j < k} \sum_k^n \frac{g_{jk}(n_i)}{g_{jk}}. \tag{3}$$

Eigenvector centrality of a node n_i , denoted by $C_E(n_i)$, is a variant of degree centrality in which a node is more central if it is connected to nodes that are themselves well-connected (Abbasi et al., 2011; Bonacich, 1972). It is computed by solving:

$$A \times c = \lambda \times c, \tag{4}$$

where A is the adjacency matrix for a graph in which $a_{ij} = 1$ if vertex i is connected to vertex j , and $a_{ij} = 0$ otherwise, c is a vector of the degree centralities for each vertex as indicated by $c = (C_D(n_1), C_D(n_2), \dots, C_D(n_n))'$, and λ is a scalar. The above equation is the characteristic equation to find the eigensystem of a matrix A (Wasserman & Faust, 1994). Then, the elements of eigenvector are the eigenvector centralities, $C_E(n_i)$, for each vertex of the graph. By convention, eigenvector centrality is given by the eigenvector with the largest eigenvalue λ (Borgatti, Everett, & Freeman, 2002). The normalized eigenvector centrality, $C'_E(n_i)$ can be found by “the square root of one half, which is the maximum score attainable in any graph” (Abbasi et al., 2011; Borgatti & Everett, 1997). Then, $C'_E(n_i)$ is given by:

$$C'_E(n_i) = \frac{C_E(n_i)}{\sqrt{2}}, \tag{5}$$

Average tie strength of a node n_i , denoted by ATS , is the proportion of the sum of unique weighted edges (the strength of a tie or an edge as the weight of the edge) that are connected to node n_i to the number of unique edges connected to node n_i (i.e., degree centrality of the node, $C_D(n_i)$). Then, similar to the calculation in Abbasi et al. (2011), for the network of collaborative outputs, ATS is calculated; by dividing a researcher's total number of collaborative outputs, NCO , with other researchers by the number of his/her reported collaborators. For the network of communication, it is calculated by dividing a researcher's total conversational exchange frequencies with other researchers, TF , by the number of his/her reported conversational partners. Then, the average tie strength is given by:

$$ATS(n_i) = \frac{\sum_k^n NCO_{ik}}{C_D(n_i)} \text{ or } \frac{\sum_k^n TF_{ik}}{C_D(n_i)}. \tag{6}$$

Efficiency coefficient proposed by Burt (1992) considers the redundancy of an individual's contacts (Borgatti, 1997). The theory of structural holes claims that the case that an individual (or ego) is connected to an individual who is in a

close-knit group is more advantageous than the case that an individual is connected to several individuals who are in the same close-knit group (Borgatti, 1997; Burt, 1992). The main reason for this is that the connections to several individuals in the close-knit group creates redundancy to the ego since information benefits provided by an individual in the close-knit group are redundant with benefits provided by other individual in the close-knit group (Burt, 1992). Burt's efficiency coefficient for non-valued and undirected relations is given by:

$$Ef(n_i) = \frac{\sum_j m_{ij} [1 - \sum_q p_{iq} m_{jq}]}{\sum_j z_{ij}}, \quad (7)$$

where p_{iq} is the proportion of node i 's network time and energy invested in the relationship with node q (node i 's contact) and calculated by

$$p_{iq} = \frac{z_{iq}}{\sum_j z_{ij}}, \quad i \neq j, \quad (8)$$

where z_{iq} is the strength of the relationship between node i and q (in binary case, 1), and $\sum_j z_{ij}$ is the total strength of the relationship with j contacts (Borgatti, 1997; Burt, 1992). m_{jq} is the marginal strength of contact j 's relation with contact q and calculated by

$$m_{jq} = \frac{z_{jq}}{\max_k z_{jk}}, \quad j \neq k \quad (9)$$

where $\max_k z_{jk}$ is the largest of j 's relations with anyone, and z_{jq} is the strength of the relations from j to q (Borgatti, 1997; Burt, 1992). Since $\max_k z_{jk}$ is 1 in non-valued and undirected graph, it becomes $m_{jq} = z_{jq}$ (Borgatti, 1997; Burt, 1992).

Local clustering coefficient: This study also considers the local clustering coefficient which is an individual's tendency toward the dense local neighborhoods. The local clustering coefficient is also defined as a measure of degree to which an individual is embedded in a tightly knit groups, i.e., positioned in a dense-connected cluster (Girvan & Newman, 2002; Hanneman & Riddle, 2005). It is necessary to consider the local clustering coefficient of a researcher because it is more likely that working in a team (or being in dense-connected cluster) leads to higher number of citations (Aksnes, 2003; Wuchty, Jones, & Uzzi, 2007). Therefore, we test the impact of the researchers' tendency toward the dense local neighborhoods on their citation performance (h-index). The Local clustering coefficient, LCC_i , of vertex i from vertices n is computed by dividing the number of edges among the neighbors of vertex i by maximum possible edges of the neighbors of vertex i (Watts & Strogatz, 1998). In other words, the local clustering coefficient calculates the density of an ego's neighbors, but by leaving out the ego (Hanneman & Riddle, 2005). Clustering coefficient for whole network, CC , is found by averaging the local clustering coefficients of all vertices n (Watts & Strogatz, 1998). That is, both of them calculated as:

$$CC = \frac{1}{n} \sum_{i=1}^n LCC_i, \quad (10)$$

$$LCC_i = \frac{\text{Number of edges among neighbors of vertex } i}{\text{Maximum possible edges of neighbors of vertex } i}. \quad (11)$$

The impact of social network metrics on the performance of individuals can be found in many studies using different types of communication and collaborative networks, e.g., the positive impact of closeness centrality in the communication network of M.B.A. students on their grade performances (Baldwin, Bedell, & Johnson, 1997), the positive impact of betweenness centrality in both friendship network and workflow network of employees in a small high-technology company on their workplace performance (Mehra, Kilduff, & Brass, 2001), the positive impact of degree centrality and network density in the advice network of employees in 5 different organizations on individual job performance and group performance (Sparrowe, Liden, Wayne, & Kraimer, 2001), and the positive impact of eigenvector centrality of group leaders in their friendship networks in the sales division of a financial services firm on the performance of their groups (Mehra, Dixon, Brass, & Robertson, 2006). Then, based on the definition of social network metrics discussed so far, the following hypotheses about the impact of a researcher's position on his/her performance are tested for each network, namely the communication network, the network of joint publications, grant proposals, and patents.

Hypothesis 1–7: The network metrics in terms of researchers' degree centrality (1), closeness centrality (2), betweenness centrality (3), eigenvector centrality (4), average tie strength (5), efficiency coefficient (6), and local clustering coefficient (7) positively impact the citation performance (e.g., h-index)

3. Method

3.1. Sample and questionnaire

The University of South Florida's College of Engineering has researchers who hold both tenured and tenure-track faculty positions, research associates, visiting professors, and graduate students to run the research. Our study surveyed the entire population, which was comprised of 107 researchers who hold both tenured and tenure-track faculty positions. Research

Table 2
Number of researchers in each demographic attribute.

	Gender		Total			
	Male	Female				
Sample	86	14	100			
Participants	68	8	76			

	Race				Total	
	Asian	Black	Hispanic	White		
Sample	35	4	9	52	100	
Participants	28	3	5	40	76	

	Department						Total
	CBE	CEE	CSE	EE	IMSE	ME	
Sample	16	19	17	24	10	14	100
Participants	14	13	10	17	10	12	76

associates, visiting professors, and graduate students were not considered in this study. The dean of the College of Engineering, 1 researcher who was on leave of absence during the data collection period, and 5 researchers who were recently hired, totaling 7 researchers, were excluded. Therefore, the sample size was reduced to 100 researchers. Table 2 shows the breakdown of the sample size in terms of demographic attributes. There are 6 departments in the College of Engineering: Chemical and Biomedical Engineering (CBE), Civil and Environmental Engineering (CEE), Computer Science and Engineering (CSE), Electrical Engineering (EE), Industrial and Management Systems Engineering (IMSE), and Mechanical Engineering (ME).

The questionnaire was in the paper-and-pencil format. It was first designed in a web format (<http://orisurvey.eng.usf.edu/>). However, several researchers during the pilot test or others later commented that filling out the questionnaire in a paper-and-pencil format was easier and more comfortable. Before distributing the questionnaire to all researchers, a researcher from each department was randomly chosen and contacted to conduct a pilot test for the questionnaire. Based on the comments and feedback from the researchers, the content and layout of the questionnaire were updated to facilitate gathering the responses.

The questionnaire was 2 pages long and contained a total of 4 questions (see Appendix). The first page included 2 questions and respondents were asked to make a self-report of the number of both in-progress and completed collaborative outputs with other researchers with whom they engaged in co-authored or joint publications (in-preparation, [re]submitted or rejected, and published), joint grant proposals (in-preparation, declined, and funded), and joint patents (rejected, submitted, and issued) as well as researchers' names. The names of the researchers from 6 different departments within the college were already populated in 6 different tables in order to facilitate the thought process of the respondents. Each table had a different number of rows due to the different number of researchers in each department and 5 columns. The first 2 columns contained the last name and first name information of the researchers populated for each department. The third, fourth and fifth columns were the columns into which the respondent put the number of total in-progress and completed joint publications, grant proposals, and patents with other researchers. Since it might be hard for the respondents to remember the exact number of their total in-progress and completed collaborative outputs with other researchers, an ordinal scale was used to facilitate the thought process of the respondents. In the scale, the scores 1, 2, 3, and 4 were assigned to the number of collaborative outputs of 1–2, 3–5, 6–9, and 10–above, respectively. For example, if a respondent has either 1 or 2 joint publications with another researcher the respondent scans the names in the tables and puts the score 1 into the related cell next to the researcher's name under the publication column. If a respondent has 3, 4, or 5 joint grant proposals with another researcher the respondent finds the his/her collaborator's name in the tables and put the score 2 into the related cell next to the researcher's name under the grant proposal column.

The second page included 2 questions and respondents were first asked to report the names of researchers with whom they exchanged conversations or ideas as well as the frequency of the exchange. The frequency in communication relations was assessed by a second question: 'How frequently do you exchange conversations or ideas?' and was rated based on a 6-point Likert-type scale (see Appendix). The second question refers to the 'frequency' dimension of tie strength¹ in the social network literature. The second page was the same as the first page except that there was one column kept for reporting the frequency of communication next to the columns across which the researchers' names were populated. Moreover, the respondent follows the same procedure which was followed to fill out the questionnaire on the first page. For example, a researcher scanned the names in the table, found his/her conversational partner's name, and put a score for the frequency

¹ Tie strength can be assessed by three indicators: the frequency of conversational exchange, the intensity of the conversational exchange, mutual confiding or level of intimacy between conversational partners (Granovetter, 1973; Marsden & Campbell, 1984).

Table 3
Timeline of the steps performed during the data collection.

Timeline	Steps
During the first week of October, 2012	A pilot test conducted for the questionnaire
In the middle of October, 2012	A mass e-mail from the dean's office was sent out to inform the researchers
During the last two weeks of October, 2012	Questionnaires began to be distributed either in the departmental meetings or through in-person delivery and e-mail
During the first week of November, 2012	A follow-up e-mail was sent to collect the completed questionnaires. The response rate was very low. Therefore, questionnaires were delivered to the researchers in person intensively. An extra 1 week was given to the participants for uncompleted questionnaires
During the second week of November, 2012	Completed questionnaires continued to be collected, and also the questionnaires continued to be delivered in person
During the last week of November and December, 2012	Due to the holiday season, there was minimum response received from the researchers
In the first week of March, 2013	All responses from the participants were finalized

of communication into the cell next to the researcher's name in a given scale. Information for the relations of both the communication (i.e., conversational exchange) and collaborative outputs between researchers was asked for the last 6 years up to current study date (between 2006 and 2012). This length of time might be reasonable for reporting the relations of the collaborative outputs, but not of communication because two researchers, for example, talk to each other frequently while they write a journal or proposal, but when they finish writing the journal or proposal they do not talk as frequently as they talked in the past. However, the main point was to investigate to what extent the researchers were genuinely acquainted with one another on average from the self-perception perspective. In addition, the time frame, 6 years, must be the same to maintain a balanced comparison between networks constructed from the relations of both the communication and collaborative outputs.

3.2. Data collection

The researchers were asked to complete a two-page questionnaire in three steps. First, a mass e-mail from the dean's office was sent out to the researchers in the sample, indicating that each of the researchers would be contacted through either their affiliated department or Second, a graduate student from the college of engineering contacted the researchers by either joining their departmental meetings or e-mailing each researcher. The student handed out the paper-and-pencil questionnaire to each researcher in the meeting and made a short presentation about the details of the questionnaire. Additionally, the questionnaire was e-mailed to the researchers who were not present in the meetings as an attachment. Last, the graduate student followed up with each researcher in the sample in 2–3 weeks for completed questionnaires via Completed questionnaires collected from the participants by visiting them directly to protect the confidentiality of their responses. If the questionnaire was not completed yet, an additional 1 week was given to the participants for completion before collecting the questionnaires directly from the participants.

Response rates were very low at the end because the number of both fully and partially completed questionnaires received was about 10. Therefore, to increase response rates, each researcher was also contacted personally both to make an in-person delivery of the questionnaire and to explain the purpose of the study and the details. The researchers were requested to fill out the questionnaire without using any forceful action which was against the protocol guidelines in the informed consent. Dillman (2007) discussed the factors improving response rate which can be achieved by in-person delivery. We observed two of those in this study. First, a deliberate effort was made to increase the salience of the experience of receiving the questionnaire; thus, the interaction time required for presenting the questionnaire to the researcher was lengthened. Second, responsibility was assigned to a researcher rather than addressing the request in a general way.

Contacting the researchers personally was performed in two steps. First, the graduate student contacted the researchers personally to deliver the questionnaire in person, explained the details of the paper-and-pencil questionnaire face-to-face, and asked for whether they were willing to participate in the questionnaire or not. Later, the researchers who were willing to participate either filled out the questionnaire at the time they were contacted personally or made an appointment with the graduate student to fill later or filled on their own. The presence of graduate student was helpful because the researchers asked if they had any questions. The questionnaire was completed in 15–20 min on average; however, a few researchers took more time to complete the questionnaire. A total of 76 out of 100 tenured/tenure-track faculty members participated in the questionnaire. Table 2 shows the breakdown of the participants in terms of demographic attributes. It took almost one semester to reach out to our target faculty members and to finalize all responses from the participants. Table 3 shows the timeline of the steps taken. One potential risk in this study was the low participation rate while collecting the social network data of researchers. If the participation rate is low, it is difficult to entirely depict connections between researchers, opening up the possibility that the results found in the analyses of the networks will be misleading. However, even if a particular faculty member did not fill out the questionnaire, the connections to non-participants are reported by the participants. Thus, connections of non-participants can be obtained from the perspective of participants. At the end, collaboration information for the full list of researchers is obtained. In this study, information about the connections of 24 non-participants was obtained by utilizing the best possible scenario explained in the next section.

Table 4
Five possible cases of reciprocity.

Cases	Upper triangle cells	Lower triangle cells
1	Equal	Equal
2a	High	Low
2b	Low	High
3a	X	0
3b	0	X

3.3. Constructing social network data matrixes

This study focuses on the population of research faculty within the University of South Florida's College of Engineering. Data was collected by employing a questionnaire by which researchers report their contacts, the number of collaborative outputs, and the frequency of communication with them in a self-reported manner. The relational data obtained through the questionnaire was put into the form of a two-way matrix where rows and columns referred to researchers making up the pairs (Wasserman & Faust, 1994). Furthermore, each cell in the matrix indicated the collaborative output or communication ties between the researchers. Thus, four 100×100 matrixes were constructed from the relational data provided by the researchers: a matrix of communication relations and a matrix of joint publications (or co-authorship), grant proposals, and patents.

Five possible cases of reciprocity happened between two researchers when they rated each other regarding their connections:

- Both researchers rated each other with an equal score for the frequency of communication and the number of collaborative outputs. In other words, the case was that the values of the upper and lower triangle cells were equal to each other in the 100×100 matrixes.
- Both researchers rated each other with a different score for the frequency of communication and the number of collaborative outputs. In this situation, two cases might happen.
 - One case was that the value of the upper triangle cells was *higher* than the value of the lower triangle cells in the 100×100 matrixes.
 - The other was that the value of the lower triangle cells was also *higher* than the value of the upper triangle cells in the 100×100 matrixes.
- Only one of the researchers rated the other. In this situation, two cases might also happen.
 - One case was that the upper triangle cell contained a value, but lower triangle cell did not in the 100×100 matrixes.
 - The other was that the lower triangle cell contained a value, but the upper triangle cell did not in the 100×100 matrixes.

Table 4 summarizes the five possible cases of reciprocity seen in the 100×100 matrixes when at least one researcher in a pair gives a non-zero rating to the other. 'X' and '0' indicate the ratings happening on only one side and non-ratings, respectively. Table 5 illustrates the number of occurrences of these cases in each network. The inter-rater agreement (IRA) percentage in a network was calculated by dividing the total number of occurrences in 'Equal-Equal' cases by the total number of occurrences of all cases (e.g., 120 was divided by 1234 which is the sum of 120, 141, 144, 377, and 452 for the network of communication). In IRA percentage calculation, we neglected the cases where both sides did not report a tie to the other, i.e., the cases where both sides score 0. For the purpose of this study, directionality of the networks is not of fundamental importance (Pepe, 2011). This is because the collaborative output networks such as co-authorship networks are analyzed as undirected in the literature. Therefore, reported reciprocity in the number of collaborative outputs was

Table 5
The number of occurrences of five possible cases in each network and inter-rater agreement percentage.

Cases	Network of communication	Network of joint publications	Network of joint grant proposals	Network of joint patents
1	120	38	81	9
2a	141	14	20	2
2b	144	16	21	2
3a	377	68	113	11
3b	452	60	132	11
Inter-rater agreement percentage	9.72%	19.39%	22.07%	25.71%

1 – The value of the upper and the lower triangle cells were equal.

2a – The value of the upper triangle cells was higher than the value of the lower triangle cells.

2b – The value of the lower triangle cells was higher than the value of the upper triangle cells.

3a – The upper triangle cells contained a value, but lower triangle cells did not.

3b – The lower triangle cells contained a value, but the upper triangle cells did not.

Table 6

The most idealistic scenario of the conversion to undirected edges.

Cases	Upper triangle cells	Lower triangle cells
1	Equal	Equal
2a*	High	High
2b*	High	High
3a*	X	X
3b*	X	X

converted to undirected edges. In order to make an equivalent comparison between the networks, the reported reciprocity in the frequency of communication was also converted to undirected edges. We symmetrized the researchers' social network data matrixes by converting the reported reciprocities to the undirected edges according to the most idealistic scenario shown in Table 6. In social network analysis, this symmetrization principle is known as the “maximum” method (Borgatti et al., 2002). All networks constructed from corresponding data matrixes were depicted in Fig. 1.

3.4. Constructing datasets for statistical models

Four datasets from four social network data matrixes corresponding to researchers' each network (e.g., communication, joint publications, joint grant proposals, and joint patents) were constructed. Each of four datasets included 11 variables for 100 researchers. In other words, four data matrixes in 100×11 dimensions were compiled. The variables included in the datasets are the researchers' citation-based performance index (h-index), 7 social network metrics obtained from each network (i.e., degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, average tie strength, Burt's efficiency coefficient, and local clustering coefficient), and 3 demographic attributes (i.e., gender, race, and department affiliation). The researchers' citation-based performance index (h-index) can be easily obtained through the Thomson ISI Web of Science database without the need for further calculation (Bornmann & Daniel, 2007). The database was accessed via the library of the University of South Florida. Each researcher's h-index was obtained by plugging the researcher's name, an organization name (e.g., the University of South Florida), and the years between 2006 and 2012 into the search boxes. The social network metrics for each network were computed using UCINET 6.308 (Borgatti et al., 2002) While centrality metrics, Burt's efficiency coefficient, and local clustering coefficient were computed using dichotomized data matrixes, average tie strength were computed using valued data matrixes.

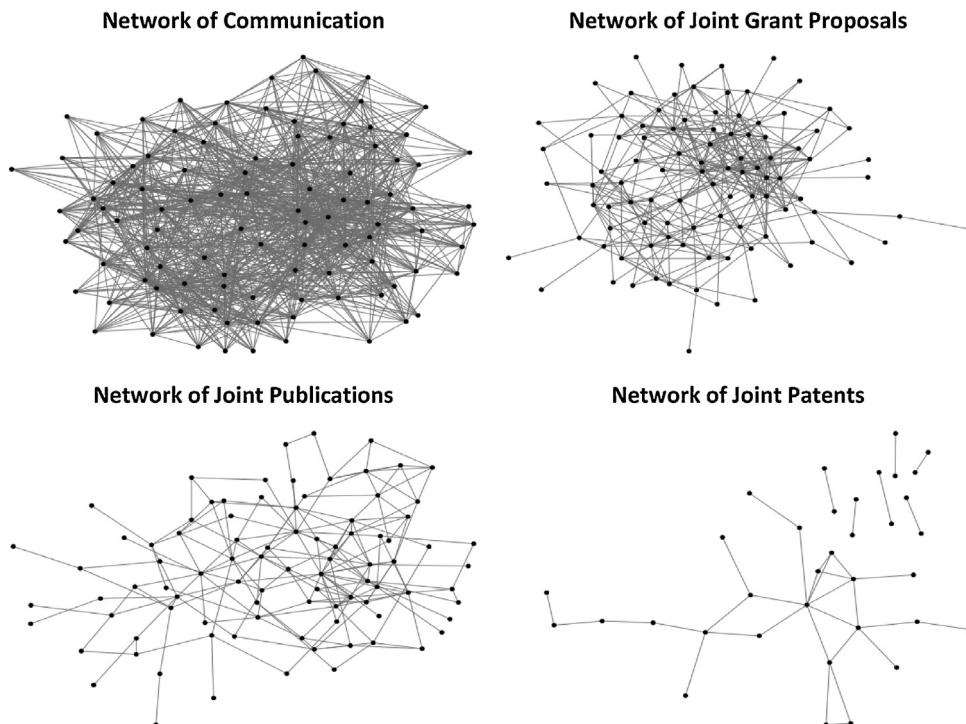


Fig. 1. Visualization of researchers' communication and collaborative output networks.

Table 7
Spearman's rank correlations.

	h-Index	C _D	C _C	C _B	C _E	ATS	Ef	LCC
<i>Communication</i>								
h-Index	1.000	0.175	0.182	0.127	0.023*	-0.055	0.033	-0.042
C _D		1.000	0.994**	0.968**	0.975**	0.033	0.855**	-0.884**
C _C			1.000	0.956**	0.985**	0.036	0.840**	-0.870**
C _B				1.000	0.912**	-0.019	0.925**	-0.945**
C _E					1.000	0.057	0.776**	-0.809**
ATS						1.000	-0.075	0.065
Ef							1.000	-0.997**
CC								1.000
<i>Joint publications</i>								
h-Index	1.000	0.422**	0.428**	0.241*	0.490**	0.456**	-0.141	0.504**
C _D		1.000	0.912**	0.835**	0.866**	0.393**	-0.124	0.616**
C _C			1.000	0.806**	0.937**	0.402**	-0.092	0.584**
C _B				1.000	0.648**	0.275**	0.110	0.281**
C _E					1.000	0.422**	-0.185	0.685**
ATS						1.000	0.096	0.363**
Ef							1.000	-0.663**
CC								1.000
<i>Joint grant proposals</i>								
h-Index	1.000	0.309**	0.316**	0.216*	0.336**	0.281**	-0.281**	0.309**
C _D		1.000	0.968**	0.847**	0.875**	0.266**	-0.249*	0.323**
C _C			1.000	0.822**	0.933**	0.267**	-0.237*	0.319**
C _B				1.000	0.664**	0.185	0.086	-0.021
C _E					1.000	0.244*	-0.334**	0.431**
ATS						1.000	-0.173	0.317
Ef							1.000	-0.840**
CC								1.000
<i>Joint patents</i>								
h-Index	1.000	0.288**	0.281**	0.077	-0.033	0.302**	0.304**	0.159
C _D		1.000	0.994**	0.641**	0.622**	0.973**	0.932**	0.532**
C _C			1.000	0.635**	0.658**	0.965**	0.930**	0.523**
C _B				1.000	0.586**	0.483**	0.474**	0.335**
C _E					1.000	0.541**	0.462**	0.511**
ATS						1.000	0.941**	0.517**
Ef							1.000	0.230*
CC								1.000

* <0.05.

** <0.01.

C_D – degree centrality, C_C – closeness centrality, C_B – betweenness centrality, C_E – eigenvector centrality, ATS – average tie strength, Ef – Burt's efficiency coefficient, LCC – local clustering coefficient.

3.5. Statistical models

Poisson regression is one of the standard (or base) count response regression models (Hilbe, 2011). The Poisson regression models were run in this study because h-index is count data, and the mean and variance of the variable h-index was reasonably close to each other (mean = 3.47 and variance = 2.78). The multicollinearity problem occurs when there is a high correlation among two or more of the independent variables in a multiple regression, meaning that one independent variable or predictor can be predicted from others (Tabachnick and Fidell, 2007). This problem can be even more explicit when social network metrics are used as predictors. The Spearman's rank correlations in Table 7 indicate that many of social network metrics, especially centrality metrics, are extremely correlated. Running a multiple regression with these highly correlated social network metrics as predictors gives unreliable estimates about an individual predictor. To overcome the challenge of potential multicollinearity between predictors, we have run a separate Poisson regression bivariate model for each of 7 social network metrics obtained from each network. The models that were run for different social network metrics in each network can be shown by:

$$\log(\text{h-index}) = \beta_0 + \beta_1(\text{a SNA metric}) + \beta_2 \text{ gender} + \beta_3 \text{ race} + \beta_4 \text{ department} \tag{12}$$

Analysis for the models was performed using IBM SPSS Statistics for Windows, Version 21.0. (Armonk, NY: IBM Corporation). In the abovementioned model, one unit increase in independent variable multiplies the mean by a factor of $\exp(\beta_j)$ (Rodríguez, 2007). The main reason for log transformation is to keep the left hand side of the equation that indicates an expected count non-negative (Cameron & Trivedi, 1998).

4. Results

Table 8 illustrates the bivariate model results for each network. Maximum likelihood estimation was used to estimate the regression coefficients of predictors (or parameters) in the model. Likelihood ratio chi-square test (also called omnibus test or test against the intercept-only model) evaluates whether or not all of the estimated coefficients are equal to zero; in other words, it is the test of the model as whole (UCLA, 2007). From the p -values, all models were statistically significant at the significance level of 0.05.

The estimated regression coefficients for each network parameter indicated the following results. Degree centrality (C_D) was statistically significant and had a positive impact in all networks except the communication network. Unlike the results of Abbasi et al. (2011), closeness centrality (C_C) and eigenvector centrality (C_E) were statistically significant, and had a positive impact on the citation performance for all networks. Betweenness centrality (C_B) had a positive significant impact for only the network of joint publications. Average tie strength (AVT) was statistically significant, and had a positive impact for only the network of joint publications and patents. Efficiency coefficient (Ef) had a positive significant impact for only the network of patents. The local clustering coefficient (LCC) was statistically significant and had a positive impact for only the network of joint publications and grant proposals.

The Poisson regression coefficients are interpreted as follows: “for a one unit change in the predictor variable, the difference in the logs of expected counts is expected to change by the respective regression coefficient, given the other predictor variables in the model are held constant.” (UCLA, 2007). For example, if a researcher in the College of Engineering increases his/her eigenvector centrality score (i.e., increase his/her connections with the researchers who are well connected) by one point in the network of communication, joint publications, joint grant proposals, and joint patents, the difference in the logs of expected h-index is expected to increase by a factor of 3.345, 3.212, 2.956, and 1.306, respectively, while the other variables are held constant in the model. The coefficients can also be exponentiated to assess the relationship between the response and predictors as incidence rate ratios (IRR) (Hilbe, 2011). For one unit increase in eigenvector centrality scores in the network of communication, joint publications, joint grant proposals, and joint patents, the expected h-index increases by a factor of 27.37, 23.83, 18.21, and 2.69, respectively (calculated as $e^{(3.345)} - 1$, $e^{(3.212)} - 1$, $e^{(2.956)} - 1$, and $e^{(1.306)} - 1$), with the remaining predictor values held constant. That is, it would be expected that a researcher with higher eigenvector centrality score in all networks has a higher h-index score than the other researchers in the College of Engineering. This result was different from the results of Abbasi et al. (2011) which found out that eigenvector centrality had a negative impact on the researcher’s citation performance. One reason for this was that the researcher was connected to other researchers who were directly connected to many individual students who already had low collaboration records. However, our results showed that a researcher can be more impactful when the researcher communicates and collaborates with other researchers who are themselves well connected. Abbasi et al. (2011) reported that including demographic information could be useful as moderating variables in the model. Since the log of expected value is modeled as dependent variable in the Poisson regression, coefficients represent the difference in the log of expected value on one level compared with another level for binary or categorical predictors (e.g., demographic attributes) (Hilbe, 2011). In almost all models, the difference in the log of the expected h-index were 0.35–0.59 units lower for females than for males, with the rest of the predictor values held constant. That is, females are expected to have 29.6–55.4% lower h-index than males are in engineering field (calculated as $1 - e^{(-0.35)}$ and $1 - e^{(-0.59)}$). For other demographic variables such as race and department, we did not observe any overall significant effects on the researchers’ citation performance.

Based on the results we have found, *hypothesis 1* is only valid when the social network metrics are obtained from the researchers’ collaborative output networks, meaning that the citation performance of a researcher improves to the extent to which the researchers have more distinct connections to other researchers in collaborative output networks than in their communication network. *Hypotheses 2* and *4* can be accepted for all networks. Then, it can be stated that an increase in occupying a central position in both communication and collaborative output networks in terms of the shortness of a researcher’s total distance to all other researchers and a researcher’s tendency to connect with other researchers who are themselves well-connected will be more advantageous to improve a researcher’s citation performance. *Hypothesis 3* only holds for the network of joint publications. This indicates that the citation performance of a researcher improves when the researcher is in the position to broker information and ideas in joint publication relations. *Hypothesis 5* can only be accepted for the networks of joint publications and patents. This means that the citation performance of a researcher improves if there is an increase in the researcher’s average number of repeated publications and patents in collaboration with other researchers. *Hypothesis 6* only holds for the network of joint patents. This means that an increasing redundancy of a researcher’s joint patent connections to a group of researchers (i.e., inventors in this case) who already generate joint patents together will improve the citation performance of the researcher. *Hypothesis 7* is only valid for the network of joint publications and grant proposals, indicating that a researcher’s increasing tendency toward the tight-knit collaborating teams when making publications and submitting grant proposals will improve the researcher’s citation performance.

5. Discussion

This study is an extension of the study of Abbasi et al. (2011), and it is performed using a richer dataset. Unlike the previous study, this study considers researchers’ social network metrics obtained from researchers’ multiple collaborative output networks constructed by self-reported data as well as social network metrics obtained from researchers’ communication

Table 8
Poisson regression results (the h-index as dependent variable) for bivariate models.

Parameter	Communication Coefficient						Joint publications Coefficient							
	Intercept	1.051 ^a	-0.051	1.316 ^a	0.864 ^a	1.632 ^a	0.884 ^a	1.769 ^a	0.885 ^a	0.063	1.210 ^a	0.930 ^a	0.734 ^a	1.154 ^a
C _D	1.077							10.232 ^a						
C _C		2.215 ^a							4.554 ^a					
C _B			1.391							5.765 ^a				
C _E				3.345 ^a							3.212 ^a			
ATS					-0.097							0.336 ^a		
E _f						0.881							0.218	
LCC							-0.838							1.200 ^a
Gender [0]	-0.441 ^a	-0.444 ^a	-0.440 ^a	-0.431 ^a	-0.430 ^a	-0.465 ^a	-0.462 ^a	-0.330	-0.293	-0.385 ^a	-0.329	-0.352 ^a	-0.416 ^a	-0.590 ^a
Gender [1]	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a
Race [1]	0.187	0.180	0.206	0.170	0.197	0.177	0.171	-0.071	-0.036	0.021	-0.104	0.108	0.194	0.167
Race [2]	-0.531	-0.550	-0.509	-0.559	-0.539	-0.592	-0.604	-0.619	-0.471	-0.539	-0.776 ^a	-0.335	-0.457	-0.390
Race [3]	-0.285	-0.301	-0.256	-0.332	-0.281	-0.301	-0.311	-0.425 ^a	-0.411 ^a	-0.433	-0.383 ^a	-0.138	-0.254	-0.168
Race [4]	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a
Department [1]	0.001	-0.004	0.051	-0.031	0.048	0.071	0.067	0.115	0.092	0.157	0.027	0.175	0.100	0.039
Department [2]	-0.126	-0.107	-0.181	-0.049	-0.238	-0.144	-0.139	0.026	-0.077	-0.044	0.184	-0.110	-0.190	-0.300
Department [3]	-0.039	-0.030	-0.080	0.018	-0.081	-0.078	-0.072	0.048	-0.078	-0.042	0.229	-0.022	-0.090	-0.176
Department [4]	0.020	0.022	0.011	0.037	-0.001	0.046	0.047	0.027	-0.098	0.055	0.033	-0.003	0.007	-0.212
Department [5]	-0.624	-0.626	-0.613	-0.620 ^a	-0.588	-0.602	-0.602	-0.446	-0.643	-0.644	-0.242	-0.495	-0.651	-0.511
Department [6]	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a
(Scale)	1 ^b				1 ^b				1 ^b			1 ^b		
Likelihood ratio chi-square	33.803	35.203	28.510	39.965	29.491	32.503	33.404	68.194	75.043	44.367	75.879	54.493	29.621	60.344
df	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Sig.	0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000

Parameter	Joint grant proposals Coefficient						Joint patents Coefficient							
	Intercept	1.038 ^a	0.160	1.283 ^a	1.020 ^a	0.847 ^a	1.814 ^a	1.159 ^a	1.167 ^a	1.181 ^a	1.305 ^a	1.230 ^a	1.215 ^a	1.076 ^a
C _D	3.613 ^a							15.583 ^a						
C _C		2.759 ^a							6.273 ^a					
C _B			3.077							15.993				
C _E				2.956 ^a							1.306 ^a			
ATS					0.339							0.223 ^a		
E _f						-0.652							0.565 ^a	
LCC							0.591 ^a							0.368
Gender [0]	-0.500 ^a	-0.446 ^a	-0.445 ^a	-0.627 ^a	-0.351	-0.480 ^a	-0.397 ^a	-0.367	-0.367	-0.442 ^a	-0.371	-0.351 ^a	-0.330	-0.416
Gender [1]	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a
Race [1]	0.141	0.123	0.184	0.124	0.191	0.251	0.217	0.215	0.186	0.199	0.195	0.276	0.243	0.215
Race [2]	-0.600	-0.633	-0.503	-0.747	-0.576	-0.590	-0.689	-0.645	-0.700	-0.579	-0.533	-0.522	-0.662	-0.505
Race [3]	-0.234	-0.311	-0.247	-0.235	-0.282	-0.175	-0.205	-0.224	-0.229	-0.268	-0.213	-0.333	-0.238	-0.240
Race [4]	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a
Department [1]	-0.030	0.079	0.032	-0.197	0.067	-0.063	-0.008	0.022	-0.067	0.055	0.057	-0.106	-0.012	-0.021
Department [2]	-0.103	-0.081	-0.163	-0.0156	-0.204	-0.197	-0.200	-0.066	-0.057	-0.156	-0.098	-0.153	-0.071	-0.182
Department [3]	0.019	0.075	-0.074	0.179	-0.052	-0.133	-0.114	0.015	0.054	-0.054	0.018	-0.187	-0.085	-0.076
Department [4]	0.040	0.045	0.005	0.079	0.001	-0.002	-0.029	-0.104	-0.158	-0.024	-0.109	-0.037	-0.008	-0.017
Department [5]	-0.511	-0.484	-0.622	-0.378	-0.639	-0.526	-0.506	-0.485	-0.535 ^a	-0.579	-0.546	-0.565	-0.483	-0.610
Department [6]	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a	0 ^a
(Scale)	1 ^b				1 ^b				1 ^b			1 ^b		
Likelihood ratio chi-square	42.813	46.668	30.224	55.364	35.763	33.979	35.747	47.404	45.676	30.267	42.088	41.195	50.803	30.738
df														
Sig.	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001

Note: 0 = 'female', 1 = 'male' for Gender; 1 = 'Asian', 2 = 'Black', 3 = 'Hispanic', 4 = 'White' for race; 1 = 'CBE', 2 = 'CEE', 3 = 'CSE', 4 = 'EE', 5 = 'IMSE', 6 = 'ME' for department.

^a Set to zero because this parameter is the base value.

^b Fixed at the displayed value.

* <0.05.

network in a small-scale such as within a college. Additionally, collecting researchers' collaborative output data in a self-reported way provides some indication of whether or not a tie is important in terms of their collaborative research efforts. In other words, the self-reported way of collecting the relations in collaborative outputs permits the researchers to assess both which connection or tie is important to them according to their own perceptions and whether or not reported contact is actually involved in research. Then, the dataset used to construct researchers' collaborative output networks contains richer data since it consists of both in-progress and completed collaborative efforts. This study also considers the local clustering coefficient, i.e., an individual's tendency toward the dense local neighborhoods. It is necessary to consider the local clustering coefficient of a researcher because it is more likely that working in a team, i.e., being in dense-connected cluster leads to higher number of citations (Aksnes, 2003; Wuchty et al., 2007). In addition, this study uses h-index instead of g-index because h-index is better to use when researchers within the same field of study are compared (Bornmann & Daniel, 2009). The Poisson regression model was used because h-index is the count data, and the mean and variance of the variable h-index was reasonably close to each other. However, as remarked by an anonymous referee, the h-index is not a pure count variable, but instead a composite index calculated from the rank-frequency distribution. Therefore, there are considerations about how to statistically analyze the h-index, which should be taken into account (Baccini, Barabesi, Marcheselli, & Pratelli, 2012). The result of Poisson regression bivariate models indicated that unlike the study of the study of Abbasi et al. (2011), eigenvector centrality (i.e., being connected to well-connected researchers) positively impacted the citation performance of the researchers. One reason for this might be that the researchers' connections with students and district connections to other researchers from different colleges are excluded. Furthermore, the previous study found out that closeness and betweenness centralities in the network of joint publications did not significantly impact the citation performance of the researchers, whereas we detected that their impact was statistically significant and positive. This study has the potential to be generalized and applied other colleges and disciplines, and even the university as a whole. However, when we apply this study to other colleges and disciplines, some of these four networks disappear. For example, writing joint grant proposals in a college of business is not as common as in a college of engineering. Moreover, some colleges and disciplines such as college of education and business have a decreased tendency to issue patents, and in some disciplines such as humanities and history, single-authored papers are more valuable than co-authored papers. This study can be run for other colleges of engineering in different universities (e.g., small-sized or large-sized, research-oriented) to understand whether the impact of social network metrics on researchers' citation performance is more or less specific for the chosen sample. In the case of extending this study to the university as a whole, this study can be run for large sample sizes including interdisciplinary collaborative output ties between researchers.

Since this study aims at evaluating the extent to which social network metrics obtained from the researchers' multiple collaborative output networks as well as their communication networks predict the performance of researchers, the information obtained from this study can be used to formulate policies that improve both the collaborative and communication relationships that impact the performance of researchers. For example, when the level of prediction of eigenvector centrality on the performance of researchers is low, meaning that the researchers tend to both collaborate and communicate with other researchers that are not well connected (i.e., other researchers that are not well-performing in their collaborative activities and communications), policies could be generated, which primarily attempt to encourage the researchers to interact with other researchers who are active in their both collaborative and communication relationships.

Acknowledgments

We are thankful to an anonymous referee who provided kind and invaluable comments.

Appendices.

First page

Question 1: With whom do you collaborate for your research matters? and

Question 2: How many in-progress and completed collaborative work do you have with other researchers including:

- In-preparation, (re)submitted or rejected, and published joint publications (column 3)?
- In-preparation, declined, and funded grant proposals (column 4)?
- Rejected, submitted, and issued patent applications (column 5)?

Scale: please put (1) for 1–2, (2) for 3–5, (3) for 6–9, (4) for 10–above

Second page

Question 1: With whom do you exchange conversations or ideas via below mentioned ways?

Face-to-face conversations: (1) formal or informal group meetings and events in department, college, and even campus level, (2) hallway conversations in department and college level, (3) serving in a student's doctoral committee, (4) telephone conversations, etc.

Conversations in virtual environment: (1) e-mail exchange, (2) exchanging ideas in online social network sites (academia.edu), etc.

Question 2: How frequently do you exchange conversations or ideas?

Scale: once a day (6), once a week (5), once every two week (4), once a month (3), once every 2 months (2), once every 3 months (1).

References

- Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4), 594–607.
- Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159–170.
- Baccini, A., Barabesi, L., Marcheselli, M., & Pratelli, L. (2012). Statistical inference on the h-index with an application to top-scientist performance. *Journal of Informetrics*, 6(4), 721–728.
- Balconi, M., Breschi, S., & Lissoni, F. (2004). Networks of inventors and the role of academia: An exploration of Italian patent data. *Research Policy*, 33(1), 127.
- Baldwin, T. T., Bedell, M. D., & Johnson, J. L. (1997). The social fabric of a team-based MBA program: Network effects on student satisfaction and performance. *Academy of Management Journal*, 40(6), 1369–1397.
- Barabasi, A. L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physics A*, 311(3–4), 590–614.
- Beaver, D. D. (2001). Reflections on scientific collaboration (and its study): Past, present, and future. *Scientometrics*, 52(3), 365–377.
- Bonach, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1), 113–120.
- Borgatti, S. P. (1997). Structural holes: Unpacking Burt's redundancy measures. *Connections*, 20(1), 35–38.
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27(1), 55–71.
- Borgatti, S. P., & Everett, M. G. (1997). Network analysis of 2-mode data. *Social Networks*, 19(3), 243–269.
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet for windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.
- Borgman, C. L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36, 3–72.
- Bornmann, L., & Daniel, H. (2007). What do we know about the h index? *Journal of the American Society for Information Science and Technology*, 58(9), 1381–1385.
- Bornmann, L., & Daniel, H. D. (2009). The state of h index research is the h index the ideal way to measure research performance? *EMBO Reports*, 10(1), 2–6.
- Bornmann, L., Mutz, R., & Daniel, H. (2008). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. *Journal of the American Society for Information Science and Technology*, 59(5), 830–837.
- Bozeman, B., & Corley, E. (2004). Scientists' collaboration strategies: Implications for scientific and technical human capital. *Research Policy*, 33(4), 599–616.
- Breschi, S., & Lissoni, F. (2004). In H. F. Moed, W. Glanzel, & U. Schmoch (Eds.), *Knowledge networks from patent data* (pp. 613–643). Dordrecht: Kluwer Academic Publishers.
- Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, 9(4), 439–468.
- Bukvova, H. (2010). Studying research collaboration: A literature review. *Working Papers on Information Systems*, 10(3), 1–17.
- Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, Mass: Harvard University Press.
- Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*. Cambridge, UK: Cambridge University Press.
- Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, 1(3), 193–203.
- Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*, 57(9), 1275–1278.
- Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science*, 35(5), 703–722.
- De Solla Price, D. J., & Beaver, D. D. (1966). Collaboration in an invisible college. *American Psychologist*, 21(11), 1011–1018.
- Defazio, D., Lockett, A., & Wright, M. (2009). Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program? *Research Policy*, 38(2), 293–305.
- Dillman, D. A. (2007). *Mail and internet surveys: The tailored design method* (2nd ed.). Hoboken, NJ: Wiley.
- Duque, R. B., Ynalvez, M., Sooryamoorthy, R., Mbatia, P., Dzorgbo, D. S., & Shrum, W. (2005). Collaboration paradox: Scientific productivity, the internet, and problems of research in developing areas. *Social Studies of Science*, 35(5), 755–785.
- Edge, D. (1979). Quantitative measures of communication in science: A critical review. *History of Science*, 17(2), 102–134.
- Fox, M. F. (1983). Publication productivity among scientists. *Social Studies of Science*, 13(2), 285–305.
- Freeman, C., & Soete, L. (2009). Developing science, technology and innovation indicators: What we can learn from the past. *Research Policy*, 38(4), 583–589.
- Friedkin, N. E. (1978). University social structure and social networks among scientists. *American Journal of Sociology*, 83(6), 1444–1465.
- Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), 7821–7826.
- Glänzel, W. (2002). Co-authorship patterns and trends in the sciences (1980–1998): A bibliometric study with implications for database indexing and search strategies. *Library Trends*, 50(3), 461–473.
- Glänzel, W., & Schubert, A. (2004). In H. F. Moed, W. Glanzel, & U. Schmoch (Eds.), *Analyzing scientific networks through co-authorship* (pp. 257–276). Dordrecht: Kluwer Academic Publishers.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Hagstrom, W. O. (1975). *The scientific community*. Carbondale: Southern Illinois University Press.
- Hale, K. (2012). Collaboration in academic R&D: A decade of growth in pass-through funding. *NSF*, 12–325.
- Hanneman, R., & Riddle, M. (2005). *Introduction to social network methods*. University of California: Riverside.
- Hansen, D. L., Schneiderman, B., & Smith, M. A. (2011). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Morgan Kaufmann.
- Hara, N., Solomon, P., Kim, S., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54(10), 952–965.
- Hilbe, J. M. (2011). *Negative binomial regression* (2nd ed.). New York: Cambridge University Press.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572.
- Hirsch, J. E. (2007). Does the h-index have predictive power? *Proceedings of the National Academy of Sciences of the United States of America*, 104(149), 19193–19198.
- Hou, H., Kretschmer, H., & Zeyuan, L. (2008). The structure of scientific collaboration networks in scientometrics. *Scientometrics*, 75(2), 189–202.
- Jiang, Y. (2008). Locating active actors in the scientific collaboration communities based on interaction topology analyses. *Scientometrics*, 74(3), 471–485.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18.

- Kraut, R., & Egido, C. (1988). Patterns of contact and communication in scientific research collaboration. In *Proceedings of the 1988 ACM conference on computer-supported cooperative work* (pp. 1–12).
- Kretschmer, H. (2004). Author productivity and geodesic distance in bibliographic co-authorship networks, and visibility on the web. *Scientometrics*, 60(4), 409–420.
- LaFollette, M. C. (1992). *Stealing into print: Fraud, plagiarism, and misconduct in scientific publishing*. Berkeley: University of California Press.
- Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation*, 11(1), 3–15.
- Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673–702.
- Marsden, P. V., & Campbell, K. E. (1984). Measuring tie strength. *Social Forces*, 63(2), 482–501.
- McCarty, C., Jawitz, J. W., Hopkins, A., & Goldman, A. (2013). Predicting author h-index using characteristics of the co-author network. *Scientometrics*, 96(2), 467–483.
- Mehra, A., Dixon, A. L., Brass, D. J., & Robertson, B. (2006). The social network ties of group leaders: Implications for group performance and leader reputation. *Organization Science*, 17(1), 64–79.
- Mehra, A., Kilduff, M., & Brass, D. J. (2001). The social networks of high and low self-monitors: Implications for workplace performance. *Administrative Science Quarterly*, 46(1), 121–146.
- Melin, G. (2000). Pragmatism and self-organization: Research collaboration on the individual level. *Research Policy*, 29(1), 31–40.
- Melin, G., & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, 36(3), 363–377.
- Meyer, M., & Bhattacharya, S. (2004). Commonalities and differences between scholarly and technical collaboration: An exploration of co-invention and co-authorship analyses. *Scientometrics*, 61(3), 443–456.
- Moed, H. F., Glänzel, W., & Schmoch, U. (2004). *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems*. Dordrecht: Kluwer Academic Publishers.
- National Science Board. (2012). *Research & development, innovation, and the science and engineering workforce: A companion to science and engineering indicators 2012 (Technical No. NSB-12-03)*. Arlington, VA: National Science Foundation.
- Newman, M. E. J. (2001a). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), 404–409.
- Newman, M. E. J. (2001b). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, 64, 016132.
- Newman, M. E. J. (2001c). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, 64, 016132.
- Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-Computer Interaction*, 15(2), 139–178.
- Pepe, A. (2011). The relationship between acquaintanceship and co-authorship in scientific collaboration networks. *Journal of the American Society for Information Science and Technology*, 62(11), 2121–2132.
- Rigby, J. (2009). Comparing the scientific quality achieved by funding instruments for single grant holders and for collaborative networks within a research system: Some observations. *Scientometrics*, 78(1), 145–164.
- Rodríguez, G. (2007). *Lecture notes on generalized linear models 2013*. <http://data.princeton.edu/wws509/notes/>
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4), 581–603.
- Schleyer, T., Spallek, H., Butler, B. S., Subramanian, S., Weiss, D., Poythress, M. L., et al. (2008). Facebook for scientists: Requirements and services for optimizing how scientific collaborations are established. *Journal of Medical Internet Search*, 10(3), 46–59.
- Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology*, 41(1), 643–681.
- Sooryamoorthy, R., & Shrum, W. (2007). Does the internet promote collaboration and productivity? Evidence from the scientific community in South Africa. *Journal of Computer-Mediated Communication*, 12(2), 733–751.
- Sparrowe, R. T., Liden, R. C., Wayne, S. J., & Kraimer, M. L. (2001). Social networks and the performance of individuals and groups. *Academy of Management Journal*, 44(2), 316–325.
- Stokes, T. D., & Hartley, J. A. (1989). Coauthorship, social structure and influence within specialties. *Social Studies of Science*, 19(1), 101–125.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.
- Tijssen, R. J. W. (2004). In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Measuring and evaluating science-technology connections and interactions: Towards international statistics* (pp. 695–715). Dordrecht: Kluwer Academic Publishers.
- Introduction to SAS. (2007). *UCLA. Statistical consulting group*. Retrieved from: <http://www.ats.ucla.edu/stat/sas/notes2/>
- Van Rijnsoever, F. J., Hessels, L. K., & Vandeberg, R. (2008). A resource-based view on the interactions of university researchers. *Research Policy*, 37(8), 1255–1266.
- Vasileiadou, E. (2009). Stabilisation operationalised: Using time series analysis to understand the dynamics of research collaboration. *Journal of Informetrics*, 3(1), 36–48.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, New York: Cambridge University Press.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442.
- Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science (Washington, DC)*, 316(5827), 1036–1039.
- Ynalvez, M. A., & Shrum, W. M. (2011). Professional networks, scientific collaboration, and publication productivity in resource-constrained research institutions in a developing country. *Research Policy*, 40(2), 204–216.