Informal Intellectual Collaboration with Central Colleagues *

Co-Pierre Georg,[†] Daniel C. Opolot,[‡] and Michael E. Rose[§]

June 18, 2017

When preparing a research article, academics engage in informal intellectual collaboration by asking their colleagues for feedback. This collaboration gives rise to a social network between academics. We study whether informal intellectual collaboration with an academic who is more central in this social network results in a research article having higher scientific impact. To address the well-known reflection problem in estimating network effects, we use the assignment of discussants at NBER summer institutes as a quasi-natural experiment. We show that manuscripts discussed by a discussant with a 10% higher than average Bonacich centrality rank results in 1.4% more citations and a 5% higher probability that an article is published in a top journal. To illustrate our results, we develop a structural model in which a positive externality from intellectual collaboration implies that collaborating with a more central colleague results in larger scientific impact of the research article.

Keywords: Informal intellectual collaboration, social network, scientific impact, centrality

JEL Classification: A14, D83, G00, O33

* We wish to thank Eckhardt Bode, Stan du Plessis, Robert Gold, Ben Golub, Edward Kerby, Han Kim, Atif Mian, Claire Reicher, Peyton Young and Yves Zenou, seminar participants at Oxford University, Kiel Institute for the World Economy, DIW Berlin, Georgia Institute of Technology, Stellenbosch University, Maastricht University, and University of Cape Town, as well as conference participants at the IfW/ZBW conference on "The Future of Scholarly Communication in Economics" 2015, the 2015 Biannual Meeting of the Economic Society of South Africa, the HAS workshop "Networks, Complexity & Economic Development" 2015, the "International Symposium on Science of Science" 2016, the 2016 Africa Meeting of the Econometric Society, and the 2016 Annual Meeting of the European Economic Association for useful comments. For valuable research assistance, we also thank Jesslyn Jonathan, Khanyisa Yolanda Nogaya and Patricia Chirwa. We are grateful to Elsevier's Scopus and in particular Lucia Schombee and Alex Uchoa for providing access to the Scopus database. The views expressed in this paper do not necessarily reflect the views of Deutsche Bundesbank.

† University of Cape Town and Deutsche Bundesbank. E-Mail: co-pierre.georg@uct.ac.za. Corresponding author.

‡ University of Cape Town. E-Mail: daniel.opolot@uct.ac.za

§ University of Cape Town. E-Mail: michael.ernst.rose@gmail.com.

1 Introduction

Economists have recognized that knowledge is a factor of production. A substantial amount of theoretical and empirical work has established the positive relationship between knowledge production and economic growth.[1] Recent studies have focused on the factors that influence the process of knowledge production and especially the role of knowledge networks.[2] A knowledge network is a system consisting of individuals that are interconnected through social ties that enable and constrain their effort to acquire, transfer and create knowledge. Studying knowledge networks is important because collaboration among researchers increased substantially over the past decades.[3] This trend is strongly supported by government policies aimed at encouraging scientific collaboration.[4]

Several authors have taken steps to explain the sources of increase in collaboration (Hudson, 1996; Laband and Tollison, 2000; Goyal et al., 2006) and show that intellectual collaboration among researchers has a positive impact on their productivity (Wuchty et al., 2007; Waldinger, 2012; Ductor et al., 2014; Ductor, 2015). By intellectual collaboration we mean both formal collaboration in the form of co-authorship, and informal collaboration in form of advice and feedback on an ongoing project.[5] Besides its positive impact on productivity, understanding how the social network of intellectual collaboration affects the outcome of research projects is also relevant in light of recent trends in academia, such as the increasing competition for space in top scholarly journals (Card and DellaVigna, 2013), the increasing time lag until publication (Ellison, 2002), the increasing duration of education of researchers (Jones, 2009). The mechanisms that drive the positive relationship between collaboration and research output are, however, not clearly understood.

In this paper, we show that intellectual collaboration positively influences research output through information spillover and complementarities in research efforts. A researcher absorbs more information about the work of other researchers when she exerts more effort on her her research activities–she becomes more knowledgeable. Collaborating with a more knowledgeable researcher implies access to more accumulated knowledge and hence greater benefits from information spillover. Our notion of intellectual collaboration is akin to the notion of absorptive capacity in the literature on R&D investment by firms. This notion was first developed by Cohen and Levinthal (1989) who show that not only does R&D generate new information, it also enhances the firm's ability to assimilate and exploit existing information. The latter leads to complementarities in R&D investments through information spillovers. We argue that the same principle applies to intellectual collaboration in fundamental scientific research.

[1] See, among others, Jaffe (1989); Adams (1990); Mansfield (1991); Cohen et al. (2002).

[2] For a recent overview of the literature on knowledge networks, see Phelps et al. (2012).

[3] Powell and Grodal (2005) and Wuchty et al. (2007) provide empirical evidence for this in both knowledge intensive industries and academia.

[4] Examples include the EU-funded Innovative Training Networks and the national Spanish Consolider Program-Ingenio 2010.

[5] See Laband and Tollison (2000) for a similar characterization of the term. Economists are prominently advised to seek feedback and commentary from their colleagues on a research article prior to publication (Green et al., 2002). This form of informal intellectual collaboration is increasingly important since publishing an article in economics takes increasingly more time (Ellison, 2002; Jones, 2009).

We develop a simple model of intellectual collaboration in academia that incorporates the notion of strategic complementarity in efforts. In our model, academics collaborate in a social network of intellectual collaboration. They derive utility from the net output from research activities. That is, the total quality and quantity of their output less the opportunity cost of not working, i.e. enjoying leisure time. We control for the quality of research output by normalizing effort by the number of projects a researcher undertakes in a given period of time. To model strategic complementarity in efforts, we assume that the marginal output depends positively on a researcher's intrinsic characteristics, and the intrinsic characteristics of- and the effort exerted by her collaborators.

The existence of a term capturing complementarity in efforts implies that equilibrium allocation of effort exercised by an academic is a function of the academic's Bonacich network centrality (Ballester et al., 2006). The aggregate equilibrium effort and hence total output increases with connectivity, but only ceteris paribus. Aggregate equilibrium effort does not necessarily increase with connectivity if the growth of connectivity also implies that a researcher's number of projects increases as well. This is because the effect of positive complementarity is diluted when a collaborator works on too many projects at the same time and exercises insufficient effort per project. Overall, in equilibrium, researchers with the highest centrality are also those with the highest effort. This implies that collaborating with central colleagues leads to higher level (quantity and quality) of individual output. At the project level, it follows that collaborating with central colleagues leads to a high quality project, and hence higher scientific impact, measured by the number of citations a research article receives and by the likelihood that the article gets published in a top journal.

We test this hypothesis in the academic field of financial economics, which has the advantage of being a large and a relatively homogeneous sub-field of economics. Our main data source are the title pages of 5808 full research articles from six scholarly journals in financial economics published between 1998 and 2011. From these articles we manually construct a novel and unique data set consisting of all authors and commenters acknowledged in a research article. We construct the social network of intellectual collaboration for each year between 2000 and 2011. Every network is constructed from the articles published in a given year and the preceding two years. In the social network of informal intellectual collaboration, authors are connected to acknowledged commenters. Links in the network are weighted but not directed.[6] Edges are not directed because information spillover occurs in both directions in intellectual collaboration. The commenter gets to know about yet unpublished results, while she provides feedback to the authors.

For our identification, however, we face the problem that social networks are inherently endogenous (Manski, 1993; Graham, 2015). Authors can write high-impact articles because they are well connected, or they are well connected because they write high-impact articles. We overcome this endogeneity problem by using a quasi-natural experiment: the assignment of discussants at the summer institutes organized by the National Bureau of Economic Research

[6] For each informal intellectual collaboration we increase the weight of the link by the inverse number of authors on the article. A link can thus represent possibly many author-commenter relationships. Our results do not qualitatively depend on the exact weighting scheme, which is why we chose a particularly simple one.

(NBER). While authors can strategically choose whom to ask for feedback about their article, they do not usually control who their discussant is. This holds especially for more junior authors.

We assess the impact of discussant centrality on citation count in a negative binomial regression and on top journal publication likelihood in a logit regression and find significant support for our hypothesis. We include a large set of control variables in our estimation, including the topic of the article as well as a number of important author, discussant and bibliometric characteristics. Author and discussant characteristics include their citation stock (as a measure of how prolific a researcher is) and their seniority in the year before publication, where seniority is given by the number of years since first publication.[7] Bibliometric characteristics include the article's number of pages, its age, age squared, the number of authors, and the publication year.

We find, in our main result, that an increase in Bonacich centrality by 100 ranks from the mean is associated with an increase by about 2.55 citations for the average article–even after addressing the reflection problem and after controlling for an extensive set of author, discussant, and bibliometric characteristics. The same increase in Bonacich centrality furthermore increases the likelihood of publishing in either a top finance or a top economics journal by about 10%. Both results are economically significant. Our results are robust to a variety of alternative specifications, including different ways of constructing the network and using different network centralities.

Our paper contributes to various strands of literature. First, it contributes to the literature that studies the impact of intellectual collaboration on research output. Hollis (2001) finds that academic teamwork has positive effects on quality, length and the number of published articles. In addition, he finds that increasing co-authorship in the past (conditioning on current average co-authorship and the lifetime number of articles) increases the likelihood that an academic is prolific today. Hollis (2001) attributes this to learning that occurs in the collaborative process. A similar result is reflected upon in Ductor et al. (2014), who find that an early career academic's network of formal intellectual collaboration helps predict their future productivity. Wuchty et al. (2007) and Ductor (2015) both document increased levels of formal intellectual collaboration in academia, and show that there is a positive relationship between co-authorship and productivity. Azoulay et al. (2010) provide evidence of information spillovers and find that the loss of a superstar academic leads to a lasting $5-8\%$ average decline in the quality-adjusted publication rates of co-authors. Waldinger (2012) shows that professors' productivity drops after they lose a co-author due to dismissal. Our paper contributes to this debate by showing that the positive relationship between intellectual collaboration and research output can be explained by the relationship between information spillover and complementarities in efforts.[8]

Secondly, our paper contributes to the theoretical literature on collaboration networks in

[7]To capture non-monotonic life-cycle effects, we also include the squared seniority in our regression.

[8]Management literature also studies whether collaboration can improve the quality and economic value of knowledge produced. Examples include Singh and Fleming (2010) who show that collaboration reduces the probability of very poor outcomes while simultaneously increasing the probability of extremely successful outcomes. Girotra et al. (2010) find that hybrid team structures, in which individuals first work alone then work together, are able to generate more ideas, generate better ideas, and to better discern the quality of ideas they generate.

academia and R&D (e.g. Goyal and Moraga-González (2001), König et al. (2014) König et al. (2015)). Goyal and Moraga-González (2001) study collaboration in $R\&D$ in the presence of externalities and show that under strong market rivalry, R&D effort declines with the level of collaborative activity. In the absence of firm rivalry, however, R&D effort increases with the level of collaborative activity. Under a similar set up König et al. (2014) show that Nash equilibrium output of firms is proportional to their Katz-Bonacich centrality in R&D network. The respective optimal output choice depends on the competition intensity the firm faces in the product market. König et al. (2015) study a model similar to ours and provide conditions for existence and uniqueness of an interior equilibrium. We differ from their model along an important: we control for the quality of research output, which in turn leads to differences in equilibrium properties of the models. More generally, our model contributes to the literature of network games that identify the role of individual centrality in the network on their equilibrium behavior (e.g. Ballester et al. (2006), Bramoullé et al. (2014) and for a comprehensive survey of the literature, see Jackson and Zenou (2015)).

Finally, our paper complements a small number of papers that investigate the relationship between informal intellectual collaboration and the research process: Laband and Tollison (2000) focus on the social aspect of informal intellectual collaboration in Economics and Biology. They find that a higher number of commenters is associated with a higher citation count over seven years. But the benefit increases in the citation count–the so-called "caliber"–of the commenter (in our nomenclature: how prolific a commenter is). Unlike Laband and Tollison (2000), Brown (2005) also includes other forms of informal intellectual collaboration, such as seminar presentations. He finds that the number of acknowledged seminars is more relevant for citation count than the number of commenters. The same is true for the acceptance probability at prestigious Accounting journals. But neither of these studies take into account the network structure of the social network prevalent in Economics and its sub-fields. Oettl (2012) takes a near-network perspective by estimating the malus co-authors of very eminent life scientists experience when these eminent scientists die. The former co-authors' drop in quality-adjusted research-output amounts to 20% in this measure. Interestingly, the most important channel in Oettl (2012) is not formal, but informal intellectual collaboration. For this reason, Oettl (2012) terms this dimension "helpfulness". Our main contribution to this literature, besides being the first to study informal intellectual collaboration in an entire academic field, is that we address the reflection problem through our empirical setup. Furthermore, our theoretical model provides a conceptual framework that allows us to empirically disentangle information spillovers from strategic complementarities.

2 A Simple Model of Collaboration in Academia

The goal of this section is to provide a simple model framework for intellectual collaboration with complementarity in efforts. We model a set $N = \{1, \cdots, i, \cdots, n\}$ of researchers who engage in research to increase the quantity and quality of their output Y_i. The production process involves individual effort as well as effort of others through intellectual collaboration; that is both formal collaboration in the form of co-authorship and informal collaboration in form of receiving feedback from other researchers. In addition to responding to a request from a colleague to provide feedback on an ongoing project, informal collaboration also includes being a discussant at a conference, as well as inter-departmental collaboration through research seminars. Intellectual collaboration contributes to output Y_i by exposing an ongoing research project to other scholars who potentially give feedback. Such feedback not only improves the quality of the current project but may also give insights to ideas for new projects, and hence even increase the total number of projects.

We assume that total output is linearly decomposable into a direct contribution and an indirect contribution from complementarities. As discussed in Section 1, complementarity in efforts contributes to research output through information spillover. At the level of an individual researcher, information spillovers with co-authors forces one to increase their effort to be able to assimilate the knowledge and techniques of co-authors. The knowledge and techniques learned not only improve the quality of the paper in progress but will also be used as input to future projects. At the level of the paper, the same notion applies with contributions from informal collaborators. A direct consequence of such interactions is an increase in the overall productivity of a researcher; that is, the marginal output. As discussed in Section 1, Several papers have documented strong evidence of such externalities in collaboration. For example Ductor (2015) finds a positive effect of intellectual collaboration on individual productivity. Wuchty et al. (2007) find an increasing dominance of teams in knowledge production, and show that teams produce exceptionally high impact research compared to solo authored work. They also find that research produced in teams tends to be more cited than that by individual authors.

Formally, let e_i denote the effort of researcher i and $\mathbf{e}$ the vector of efforts. Let G be a collaboration network among researchers. With slight abuse of notation, we also write G for the adjacency matrix of G. That is, each element g_{ij} of G is defined in such a way that $g_{ij} = g_{ji} = 1$ if researcher i collaborates with j and zero otherwise. We assume that the links are undirected. For the case of informal collaboration, this results from the idea that in the process of giving feedback, a researcher also learns about the methods and results of someone else's work. Ideally, the value of g_{ij} should be different, depending on whether j is a co-author or informal collaborator. This would clearly distinguish the level of information spillover between the two types of interactions. Here, for the sake of simplicity, we assume that the level of information spillover is identical in both cases. Moreover, there is evidence suggesting that informal collaboration is just as effective as formal collaboration in generating ideas for research. For example Colander (1989, p. 146) concludes his survey that "[m]uch if not most of the debate and discussion about economic ideas take place at the pre-working paper, workshop and working paper stages.", and Ductor et al. (2014, p. 937) argue in a study on productivity patterns among co-authors that *"a researcher who is close to more productive researchers may have early access to new ideas"*. Nevertheless, in our empirical analysis below, we consider both the network of only informal and a combined network of formal and informal collaboration.

For each i, let N_i denote the set of first-order neighbors. That is, the set of all agents who directly collaborate with i. Let n_i be the cardinality of N_i. To control for quality of each output, let p_i be the total number of projects i is involved in. Assuming that i allocates her effort equally across the p_i projects, effort exerted on each project is $\frac{e_i}{p_i}$.[9] The total output $Y_i(G, \mathbf{e})$ to i for engaging in collaboration network G, given effort configuration $\mathbf{e}$ is then:[10]

$$Y_i(G, \mathbf{e}) = e_i t_i + \sum_{j \in N_i} g_{ij} \frac{e_j t_j}{p_j} + \alpha \sum_{j \in N_i} g_{ij} \frac{t_j}{p_j} e_j e_i \quad , \tag{1}$$

where t_i, the type of i, is individual productivity of i, which is determined by i's intrinsic characteristics such technical skills and seniority in the field. The first two terms on the right hand side of (1) capture the direct contribution of i and her set of collaborators. The third term captures complementarity in efforts, where α is a parameter capturing the contributive strength of such externalities. Following the above discussion, we see from (1) that the overall productivity of i is $t_i + \sum_{j \in N_i} g_{ij} \frac{e_j t_j}{p_j}$. The second term captures the notion that the characteristics and effort of i's collaborators positively influence her productivity.

Associated with effort level e_i, is an opportunity cost $c_i(e_i)$. We let $c_i(e_i)$ assume a quadratic form with parameter β identical for all agents for simplicity, i.e. $c_i(e_i) = \frac{1}{2}\beta e_i^2$. The utility $U_i(G, \mathbf{e})$ that i derives from network G while exerting effort e_i is then the net output $Y_i(G, \mathbf{e}) - c_i(e_i)$. That is:

$$U_i(G, \mathbf{e}) = e_i t_i + \sum_{j \in N_i} g_{ij} \frac{e_j t_j}{p_j} + \alpha \sum_{j \in N_i} g_{ij} \frac{t_j}{p_j} e_j e_i - \frac{1}{2}\beta e_i^2. \tag{2}$$

The model specification in (2) has similarities with those in the literature of network games (e.g. Ballester et al. (2006), Bramoullé and Kranton (2007) and Bramoullé et al. (2014)). The similarity is in the existence of externalities. The difference arises in the nature of the production process we model, which involves group production as captured by the first two terms of (2), and that we control for the quality of output. We characterize equilibrium properties of the game, and in particular conditions for existence and uniqueness of an interior equilibrium solution. We then use the resulting equilibrium efforts in an empirical study where we test for the existence of complementarities and information spillover in collaboration. We achieve this by assuming that equilibrium effort of a collaborator has a positive impact on the quality and hence impact of a researcher's output. We proxy the impact of research output by the citation count it receives. In particular, we estimate models of the following kind:

$$Citation_i = F(\mathbf{t}, \mathbf{p}, \mathbf{e}^*, \mathbf{c}), \tag{3}$$

where we write $F(a)$ to imply a function of a. The independent variables are $\mathbf{t}$ which is a vector of types, $\mathbf{p}$ the vector of number of projects, $\mathbf{e}^*$ which is equilibrium outcome and $\mathbf{c}$, a vector of control variables. We use citations of research output as a measure of its impact.

[9] This assumption simplifies our computations and comes at no loss of generality: a researcher will spend more time on projects where she is a co-author, but this could be easily adjusted for by using a simple multiplicative factor.

[10] Note that the first term of the right hand side of (1) results from summing $\frac{e_i}{p_i}$ over all projects p_i.

3 Equilibrium Properties

Given (2), each researcher chooses an optimal level of effort e_i^*, where the first order condition is:

$$\beta e_i^* = t_i + \alpha \sum_{j \in N_i} g_{ij} \frac{e_j^* t_j}{p_j} \quad \text{for each } i \in N. \tag{4}$$

It is well known (e.g. Ballester et al. (2006)) that equilibrium levels of effort in a game set up such as (2) depend on the Bonacich centrality of the underlying network of interactions, defined as follows. Given a scalar $\lambda \geq 0$ and a network G, let a matrix $M(G, \lambda)$ be defined as

$$M(G, \lambda) = (I - \lambda G)^{-1} = \sum_{k=0}^{+\infty} \lambda^k G^k$$

where I is the identity matrix. Let $\mathbf{1}$ be a column vector of ones. For an $n \times n$-square matrix G and a scalar λ such that $M(G, \lambda)$ is well defined, the vector of centralities of parameter λ in G is:

$$\mathbf{b}(G, \lambda) = (I - \lambda G)^{-1} \mathbf{1}$$

The Bonacich centrality of node i is $b_i(G, \lambda) = \sum_{j=1}^{n} m_{ij}(G, \lambda)$, and counts the total number of paths in G starting from i. For a vector $\mathbf{t}$ of t_i's, we define a corresponding vector of centralities $\mathbf{b}(G, \lambda, \mathbf{t})$ as

$$\mathbf{b}(G, \lambda, \mathbf{t}) = (I - \lambda G)^{-1} \mathbf{t}$$

For the remainder of the paper, we write $\mathbf{p}$ for a vector of p_i', D_{t_p} and D_t for the diagonal matrices consisting of t_i / p_i's and t_i as diagonal elements respectively. We also define $A = D_{t_p} G$ and $\alpha_\beta = \frac{\alpha}{\beta}$.

Proposition 1. *Let $\mu_1(G)$ be the maximum eigenvalue of G, $p_1 = \max_{i \in N} p_i$. and $t_n = \min_{i \in N} t_i$. The game with payoffs in (2) has a unique interior equilibrium whenever $\beta p_1 > \alpha t_n \mu_1(G)$. The respective equilibrium configuration is given by*

$$\mathbf{e}^* = \frac{1}{\beta} \mathbf{b}(A, \alpha_\beta, \mathbf{t}) \tag{5}$$

Proof. See Appendix A.1 □

Proposition 1 provides a characterization of equilibrium efforts. Equilibrium efforts are a function of Bonacich centralities of the network, a property that is well known in the literature of network games. There are two main differences between the measure of centrality in (5) compared to the existing literature. The first difference being that in (5), the network is weighted (inversely) by the number of projects D_p^{-1}. Note that $D_{t_p} = D_t D_p^{-1}$. At an individual level, as described by (4), being connected to another researcher engaging in many projects leads to lower combined equilibrium effort. The underlying reason is that a researcher with a large number of projects allocates relatively less effort to each of her neighbours. This in turn reduces the contribution of complementarity to the effort exerted by her set of collaborators.

When examining the effect of connectivity on equilibrium efforts at an aggregate level, the proportional change in the number of projects should also be taken into account. It is well known in the literature (e.g. Ballester et al. (2006), Bramoullé et al. (2014) and the references therein) that increasing the level of connectivity increases the overall equilibrium effort. A similar result applies as a corollary to Proposition 1 but only ceteris paribus. If creation of new links results from new projects, then equilibrium effort need not necessarily increase, at least not in the same proportion as when the number of projects stays constant. The following corollary formalizes these arguments. Consider a specific case where all researcher are engaged in the same number of projects p.

Corollary 1. *Consider two networks G and $G' = G + D$, with $p_i = p$ and $p'_i = p + x$ for all $i \in N$ as number of projects, respectively. Let $\mathbf{e}^*$ and $\mathbf{e}^{*'}$ be the respective equilibria, and S_e and $S_{e'}$ the respective sums of elements of $\mathbf{e}^*$ and $\mathbf{e}^{*'}$. If $t_i = t$ for all $i \in N$, then $S_{e'} > S_e$ if and only if*

$$x(\beta S_{ee'} - tS_e) < \alpha t \mathbf{e}^{*'} D \mathbf{e}^{*^T} \tag{6}$$

where $S_{ee'} = \sum_{i \in N} e_i^* e_i^{*'}$.

Proof. See Appendix A.2 □

Corollary 1 provides a condition under which aggregate equilibrium increases when simultaneously increasing the number of collaborators and number of projects. If the number of projects is kept constant $x = 0$, then from (6) $\mathbf{e}^{*'} D \mathbf{e}^{*^T} > 0$ and (from (18)) $S'_e = S_e + \frac{\alpha}{p} \mathbf{e}^{*'} D \mathbf{e}^{*^T}$; hence aggregate equilibrium outcome increases with the number of collaborators. For x sufficiently large however, there exists values of β and t for which aggregate equilibrium outcome need not increase.

The second component of the centrality measure in (5) that is specific to our model is its dependence on the distribution of types of researchers. researchers with higher individual productivity exert higher effort at equilibrium, as observable from (4). On an aggregate level, researcher with higher productivity have a positive impact on aggregate effort. That is, since they tend to exert a higher effort, other researcher who directly collaborate with them would also exert higher effort due to effect of complementarities, and so will the collaborators of collaborators, and so on. This effect is observable from (5) and the fact that $\mathbf{b}(A, \alpha_\beta, \mathbf{t}) = \left(I - \frac{\alpha}{\beta} A\right)^{-1} \mathbf{t}$, where the term $\left(I - \frac{\alpha}{\beta} A\right)^{-1}$ acts as a multiplier on the types. This effect is strongest if individuals with the highest productivity are also the most central.

There are three main conclusions that follow from our stylized model of intellectual collaboration. First, high connectivity , and hence high intensity of collaboration among researchers leads to higher aggregate equilibrium outcome, but only ceteris paribus. The second and related conclusion, is that increasing the number of projects per researcher may dilute the effect of positive complementarities and hence can have a negative effect on aggregate equilibrium outcome. Third, for a given network of collaboration and distribution of types, the optimal aggregate equilibrium outcome is obtained in a set up where the most central individuals are also those with the highest individual productivity. In the following, we bring these results to the data and test them empirically.

4 Data

We use data from three different sources. Our main data source is hand-collected information about informal collaboration from the acknowledgments section of 5808 articles published in six scholarly journals in finance, which we use to build the social networks of informal intellectual collaboration. To address the reflection problem of Manski (1993) we use 389 articles that have been presented at finance-related NBER summer institutes. Finally, we access Elsevier's Scopus database to obtain citation counts for each article as well as publication records for 85% of the researchers (authors and acknowledged commenters) in the network and all the authors and discussants of the published NBER articles.

We obtain the acknowledgements sections from research articles published in six scholarly journals in financial economics. We focus on journals with a similar topical focus to avoid issues that might arise from different conventions for acknowledgements in different fields. The six journals we have selected are: The Journal of Finance (JF), the Journal of Financial Economics (JFE), The Review of Financial Studies (RFS), the Journal of Financial Intermediation (JFI), the Journal of Money, Credit, & Banking (JMCB), and the Journal of Banking and Finance (JBF). The first three journals are commonly regarded as the top journals in financial economics that speak to a broader audience, which is why we denote them as general interest journals, while the other three journals are more specialized, which is why we denote them as field journals.[11] As of 2014, the general interest journals have an impact factor well above five, while the field journals' impact factor is between 1.2 and 2.5.[12]

We look at articles published in the 1998-2011 period which Scopus classifies as either article, conference paper, or review. That is, we omit notes, discussions, shorter articles, conference announcements, letters, and policymaker roundtables. This gives us a total of 5808 original articles.

The sample is very homogeneous: 92% of the 4085 articles listing Journal of Economic Literature (JEL) codes belong to general category G (Financial Economics).[13] Additional 6% list E (Macroeconomics and Monetary Economics), but not G. The six journals are not only homogeneous in their research focus, but also in their editorial process: All journals except the

[11] The set of general interest journals is identical to the set used in the study of Borokhovich et al. (2000), who examine the impact of formal collaboration in financial economics. In their annual reports the JF refers to these journals as top journals as well.

[12] Impact factors are obtained from `http://www.oxfordjournals.org/our_journals/rof/about.html`. Impact factors change over time, but the three general interest journals had a significantly higher impact factor throughout our sample period.

[13] Not all articles list JEL codes. The JF for instance never lists JEL codes, while the RFS introduced JEL codes in Winter 2006 only. For articles not listing JEL codes, we conduct an internet search to obtain JEL codes from latest working paper version. This was the case for 222 articles.

JMCB have a single blind referee process throughout the sample period. On top of that, three journals (RFS, JFE and JBF) explicitly encourage authors to acknowledge helpful individuals on the article.[14]

For each research article, we collect standard bibliometric information from the title page. This includes the title, the names of all authors, all affiliations of all authors, listed JEL codes and the number of pages.

We are mainly interested in the article's acknowledgements, typically located on the title page. Authors commonly acknowledge helpful input by colleagues and state at which conferences and seminars an article has been presented. Funding sources and hospitality during visiting positions are often acknowledged, too, as well as help from research assistants. From the acknowledgement section, we collect at which seminars and conferences the article has been presented, and, crucially, the names of the colleagues that are acknowledged for intellectual input.[15]

Cronin (1995) distinguishes three broad forms of acknowledgements: resource-related (funding, data and materials), procedure-related (editorial and moral support) and concept-related (ideas, feedback and commentary). We focus on concept-related acknowledgements, and, within this category, distinguish the following groups of acknowledged individuals: (1) editors, (2) referees, (3) discussants, (4) PhD advisers and committee members, (5) colleagues that have provided comments (commenters), (6) colleagues that have provided data, (7) research assistants, and (8) non-academic personnel from banks and industry. We restrict our analysis to categories (3) to (5) in this study. There are several reasons for this focus. First, categories (6) to (8) are not predominantly about the flow of information between academics. While research assistants often have valuable input in a research project, they are usually students and we are interested in the flow of information among colleagues. Second, the vast majority of articles acknowledges the editor of the respective journal. If we calculate an editor's position within the social network of informal collaboration, we are likely to be biased towards more frequently publishing journals. The more articles a journal publishes, the better is its editor's perceived centrality in the uncorrected data. We avoid this issue by excluding editors during their tenure from our sample.[16]

A consolidation procedure for all 13,067 names in our database is necessary because the same name is frequently spelled in different ways. The Journal of Finance's longtime editor Campbell R. Harvey, for example, is being acknowledged as Cam Harvey, Campbell Harvey, Campbell R. Harvey, and Campell Harvey (with a typo). An additional problem arises for example with different naming conventions e.g. for Asian names (first name, last name vs. last name, first name) and when family names change due to marriage. To account for all these effects, we conduct an internet search for all authors and acknowledged individuals to find their full and proper name. We then use this information to consolidate the author and commenter names.

[14]We are grateful to the journal editors for answering a short questionnaire about the editorial policies of their journals.

[15]The words that authors use to indicate input by their colleagues are usually: comments, insights, encouragements, discussions.

[16]We obtain this information either from the journals' website, or with the help of their editorial offices.

After the consolidation process, we search for Scopus Author IDs to obtain author metrics via Scopus' Author Retrival API. To obtain Scopus Author IDs for authors we use the information provided in the Scopus entry for the respective articles. The match of acknowledged commenters follows a more sophisticated procedure. The procedure is necessary because usually bibliographic databases do not aggregate perfectly. Frequent errors are too many author profiles for the same real person, and multiple real persons aggregated in one (or many) author profiles. Finally there is a beta error of obtaining a false match when otherwise the names match perfectly For this reason the search is largely performed manually but computer-aided. There are two general conditions to match a person with a Scopus Author profile: The profile is classified by Scopus as working in at least one of the fields "Economics, Econometrics and Finance" and "Business, Management and Accounting", and does not include more than 5% publications in journals outsides these fields. If only one match is found against the Scopus database via a simple name search we match the person. If the search returns up to 5 profiles satisfying above conditions, and they all list with the same name and affiliation, we take the profile with the highest publication count. In case more profiles are returned, or the returned profiles do not match in affiliation and/or name, we perform a manual search for all individuals that are acknowledged more than 5 times or are listed as discussant.

Following these procedures we match 6408 out of 6408 (100%) of the authors of articles from the core journals (i.e. that we use to construct the networks from), 629 out of 629 (100%) of the authors of the publications from the *NBER sample*, 261 out of 261 (100%) of the discussants of these publications, and 9042 out of 11857 (74.24%) of the acknowledged commenters (includes commenters that are also authors). Note that not all acknowledged commenters and discussants are listed on Scopus. In order to have a Scopus profile, an author must have published at least once in a journal or book that Scopus indexed. Regarding the computation of Bonacich centralities of acknowledged commenters in the network of informal intellectual collaboration only members of the network's largest component matter. That is, we only require nodes from the respective giant components to have verified Scopus ID. In these networks the share of nodes having a Scopus ID ranges between 85.72% (for the 2000 network) and 88.44% (for the 2005 network). Hence the error resulting from missing records (2596 distinct nodes across all the networks) is arguably small. To assess the impact of missing data, we compare results before and after adding 250 author profiles: Regression coefficients remain unchanged, but increase in precision in terms of lower p-value.

[TABLE 1 ABOUT HERE]

Throughout our paper we use three different samples which we summarize in Table 1 in Appendix A. The *full sample* is the largest sample in terms of sample size and contains all 5808 research articles for which we hand-collect the acknowledgements information. We use this sample to describe a number of stylized facts in Section 5. In Section 5.3, we consider all research articles in the *full sample* that acknowledge at least one dated conference to study the relationship between a research article's age and measures of informal collaboration (we denote this sample as the *age sample*). For our main identification in Section 6 we use the *NBER sample* where we consider only those research articles that were presented at an NBER summer institute. This is contrasted with the *commenter sample* in Section 6.3, where we consider all research articles from the *full sample* that acknowledge at least one commenter. We use this sample to contrast the exogenous (from the author perspective) attribution of discussants at NBER summer institutes with the endogenous selection of commenters.

All dependent and independent variables used in any of our regressions are summarized in Tables 2 and 3.

[TABLE 2 ABOUT HERE]

[TABLE 3 ABOUT HERE]

5 Stylized Facts about Informal Intellectual Collaboration

5.1 Do Authors Acknowledge Strategically?

Before going into any details about how academics collaborate informally, we must address the important question whether acknowledgements contain actual information about collaboration patterns, or whether authors use their acknowledgements strategically to influence editor or referee decisions. We define strategic acknowledging as an author's attempt to influence an editor in her choice of referees (Hamermesh, 1992). The assumption authors could make is that editors do not pick acknowledged commenters because commenters have already seen the article.[17] According to this view, authors would want to thank someone that has not, or only briefly, seen or heard about the article, but who has a reputation of being a tough referee.[18] Strategic acknowledging carries a high reputation risk, though. If an author acknowledges someone who has never heard about the article, there is a risk that the commenter will learn about it, which will reflect very badly on the author. In the worst case, that person might still become referee or acting editor, inflicting greater reputational damage once she sees the acknowledgment section.[19]

While we cannot rule it out entirely, we observe a number of stylized facts that provide evidence against strategic acknowledging. First, the increased competition and scrutiny for publishing in a top journal makes it less likely that signaling would be successful. Hence, research articles published in lower-ranked journals would be more likely to include a signal and therefore contain more acknowledgements. We observe the opposite pattern, however. All measures

[17] There are conflicting views about this issue: Some editors have suggested to us that editors would want to pick acknowledged individuals, while many of our colleagues believe that editors do not consider the acknowledgements at all.

[18] This strategy is summarized in *"Cite your friends, acknowledge your foes."* Editors of various journals, however, have indicated to us that they seldom exclude a potential referee simply because this person is acknowledged in the article. Also, not all editors explicitly look into the acknowledgment section when selecting a referee.

[19] Because of this risk, Hamermesh (1992, p. 171) writes in his "Guide to Professional Etiquette": *"DON'T PLAY THESE GAMES - the gains are not worth the potential cost of being caught"* (emphasis in the original).

of informal collaboration are larger at higher-ranked journals. Second, authors do not put the names of senior and prominent colleagues first in the acknowledgment section. Rather, authors usually order the names of commenters alphabetically. Sometimes editors and referees are thanked separately before all other commenters. Third, the list of commenters is not always first in the acknowledgement section, which in total may well span more than 10 lines. Seminars, conferences, research assistance or funding are listed before commenters are listed. And fourth, more than half of all articles acknowledge individuals that no other publication acknowledges. There is little signaling value in adding relatively unknown colleagues, which we take as indication that these colleagues have provided substantial input. Finally, in Georg and Rose (2015), we show that frequently acknowledged commenters are not necessarily the most important ones for the flow of information, nor the most prolific authors, which is further evidence against the presence of strategic acknowledging.

A less severe form of strategic acknowledging could still exist in our data. Authors could, for example, strategically seek advice from senior and well-known colleagues. This variant of strategic acknowledging, however, is precisely what we want to capture. Authors identify scholars that they think might be of help for an ongoing research project and to which they establish a tie. For our analysis, it is not relevant why scholars discuss with each other, as long as information actually flows.

5.2 The Extensive and Intensive Margin of Informal Collaboration

The vast majority of published research articles in our sample acknowledges informal input by colleagues. Of the 5808 articles in our dataset, 5222 (≈ 90%) articles do acknowledge at least one commenter or one seminar or one conference.[20] Overall, the extensive margin of informal collaboration is very stable over time, albeit research articles in general interest journals report informal collaboration significantly more often. Figure 1 documents this development.

FIGURE 1 ABOUT HERE

General interest publications not only acknowledge informal collaboration more often, they also report it with a higher intensity. Figure 2 shows the average number of authors, commenters, seminars and conferences, grouped by year and journal. General interest publications acknowledge almost twice as many colleagues as research articles published in a field journal (left panel in Figure 2), and are presented more than twice as often at seminars and conferences (center and right panel in Figure 2).[21]

FIGURE 2 ABOUT HERE

[20] The remaining articles may, however, acknowledge the editor, anonymous referees, funding, data exchange and research assistance.

[21] The record for the number of commenters is held by Spamann and Holger (2010): 'The "Antidirector Rights Index" Revisited', *The Review of Financial Studies* **23**(2), 467-486, which acknowledges as many as 53 individuals.

Another striking fact is how well journal impact factors are reflected in the amount of informal collaboration. The Journal of Finance usually has the highest intensity of informal collaboration, with the Journal of Banking and Finance usually having the lowest. All three general interest journals are relatively close together in terms of collaboration patterns. Not so the field journals, however, where collaboration patterns in the Journal of Financial Intermediation resemble those in general interest journals and are clearly distinct from collaboration patterns in the Journal of Banking and Finance.

Authors typically rely on all three forms of intellectual collaboration, i.e. talking to colleagues, giving seminars, and presenting at conferences. Figure 3 shows that, from 5222 articles with acknowledgements, 48% acknowledge at least one colleague, at least one seminar and at least one conference. A further 20% of all articles acknowledge only colleagues and another 16% acknowledge colleagues and seminar presentations.

FIGURE 3 ABOUT HERE

Informal collaboration is not significantly different for solo-authored work than for co-authored work. Of the 5808 articles in our database, 1258 are solo-authored and more than 91% of them (1148) report any form of informal collaboration. This compares to 89% or 4074 multi-authored articles with acknowledgments from a total of 4550 multi-authored articles in our dataset. The intensive margin for informal collaboration is equal as well: The median number of commenters acknowledged is 7, both for solo-authored and multi-authored articles. An inverse relationship in the amount of formal and informal collaboration is consistent with the view of optimal academic team sizes (Haeussler and Sauermann, 2016).

5.3 How a Research Article's Age Relates to Measures of Informal Collaboration

We estimate the age of a research research article using information from acknowledgment section, where authors often indicate which conferences they have visited and, crucially, when. We define the age of a research article as the difference between publication year and the year of the oldest conference listed in the acknowledgement section. Using this method, we are able to estimate the age for 2,314 research articles. The mean age of a research article in this sample is 2.94 years (see Table 4), with the median age being 3 years. Figure 4 visualizes the distribution of research articles' age.

TABLE 4 ABOUT HERE

FIGURE 4 ABOUT HERE

Correlation with the amount of informal collaboration, i.e. the number of commenters, seminars and conferences acknowledged is ambivalent (see Table 5): There is no strong correlation with measures of informal collaboration. That is, Pearson correlation coefficients between

Age and the number of commenters, the number of seminars and the number of conferences is 0.08, 0.16, and 0.14 respectively, while Spearman correlation coefficients are 0.09, 0.13, and 0.18.

TABLE 5 ABOUT HERE

It is hence not clear how the extent of informal collaboration in a research article at the time of publication is related to the article's age. One hypothesis is that the extent of informal collaboration increases in age, because authors have more time to polish their research article. For this to be the case, authors would have continuously keep improving the research article during the publication process. The alternative is that the extent of informal collaboration does not increase in the age of a research article, for example because authors stop asking for feedback once an article has reached maturity.

We define the amount of informal collaboration in three ways: the number of commenters, the number of seminars and the number of conferences, as acknowledged in the research article's acknowledgement section. We test this hypothesis in a negative binomial regression, because the dependent variable is a count variable and highly skewed in each of the three dependent variables (Mullahy, 1986).[22]

TABLE 6 ABOUT HERE

Table 6 reports the result of the regression at the sample mean. For each of the three regressions, we include year- and journal-fixed effects to account for time trends and journal-specific collaboration patterns. We also include the squared age to capture non-monotonic relationships. Column 1 indicates that the number of acknowledged commenters increases by 8% (or $0.08 \cdot 9.83 \approx 0.8$ commenters) for every additional year from the mean. Column 2 indicates that the number of acknowledged seminars increases by 10% or 0.5 seminars for every additional year from sample mean. Column 3 indicates that the number of acknowledged conferences increases by 22.4% or 0.63 conferences for every additional year from sample mean. All coefficients are highly statistically significant. The coefficient for squared age is highly statistically significant and negative for the conferences specification. This indicates a non-monotonic relationship for age and number of conferences, but not for age and the number of commenters or the number of seminars. This pattern is consistent with a production cycle for research articles in which authors get feedback from their colleagues and friends first, then submit to conferences, and ultimately to journals. Once papers get accepted at conferences (we only observe journal acceptances, not journal submissions), direct feedback from colleagues becomes less important since the articles have reached a certain maturity.

[22] To ease interpretation, we report marginal coefficients: The coefficients we present are the expected *percentage* increase in the outcome variable when the explanatory variable increases by 1 unit and when all other variables are held constant at their mean and all dummy variables at 0.

5.4 The Social Network of Intellectual Collaboration

Several papers have studied the role of networks in knowledge creation and sharing (See Phelps et al. (2012) for a survey). Parallel to this literature is a body of theoretical models on strategic interactions in social networks (see Jackson and Zenou (2015) for a survey of the literature). The latter literature studies equilibrium behavior in situations where agents' actions are strategic complements (and strategic substitutes). A recurring theme in this literature is that equilibrium actions depend on an agent's centrality in the network. Our paper brings together these two strands of literature by arguing that knowledge spillover that occurs through collaboration creates synergies that are modeled as positive complementarities. We test for this processes by hypothesizing that the quality of a research output depends on the centrality of those contributing to its development.

We test for existence of knowledge spillovers in intellectual collaboration by considering both a network of informal collaboration and a combined network of formal and informal collaboration. To construct a network of informal collaboration, we connect two nodes whenever one acknowledged the other as a commenter on a published research article in our dataset. A commenter is a colleague that has been listed in an article's acknowledgement section, except when a) she is the journal's managing editor, b) she was thanked for data collection or research assistance, c) she is an industry or bank employee who participated in surveys or similar, and d) she is a discussant or known referee of the article.

For each year t we construct the network using the publications published in t, as well as in the two previous years, $t-1$ and $t-2$. We construct twelve networks for all t between 2000 and 2011. As the number of articles increased over time, the network increase by size, too: The 2000 network is generated from 873 articles published in either 1998, 1999 or 2000 and consists of 3289 distinct researchers. This compares to the 2011 network, which connects 6997 researchers that have collaborated on 1888 articles.

In our social networks, all ties are directed but exist twice. This is based on the notion that knowledge spillover occurs in both directions: The author tells the commenter, sometimes in great detail, about the article, which is valuable to the commenter. For example she can use the results to build her own research on it before it is published. Knowledge spillover from the commenter to the author, occurs in the form of feedback, which in turn not only improves the quality of the author's current work but may also provide ideas for future research.

Links are weighted to reflect a) the level of interaction between collaborators and b) the productivity of the node from which the link starts. Formally, let A_t be the set of articles published in years $\{t, t-1, t-2\}$. To each article $a \in \{A_t\}$, there is a non-empty set of authors κ_a and a not necessarily non-empty set of commenters ι_a. Every author $i \in \kappa_a$ and every commenter $k \in \iota_a$ is part of the set of nodes that either authored or acknowledged in the set of articles A_t. The resulting network G is weighted in such a way that for each pair i, j, g_{ij} increases by $1/|\kappa_a|$ if author i acknowledges commenter j on article a. If the commenter has been acknowledged once on an article written by two authors, the weight of each of the two ties would be 1/2. If one of the authors acknowledges this commenter on another solo-article, the tie increases to

3/2, reflecting a deeper relationship between the two. The adjacency matrix G is symmetric as acknowledgements are undirected, with the diagonal elements being equal to 0. This weighting scheme also corrects for misreporting. That is, when a research article consists of many authors, it is not clear which author spoke to which commenter. This is thus accounted for by the weights of $1/|\kappa_a|$.

Our main regression equation is (3), which is elaborated further in Section 6 below. The dependent variable is the quality of a research output as measured by the quantity of citations it received. Among independent variables is equilibrium level of effort of scientists contributing to its development, and of particular interest is the commenters. In (5) of Proposition 1, we show that equilibrium efforts $\mathbf{e}^*$, are equal to the Bonacich centralities $\mathbf{b}(A,\alpha,\mathbf{t})$ of network A weighted by the vector $\mathbf{t}$ of individual productivities.[23] That is

$$\mathbf{e}^* = \mathbf{b}(A,\alpha,\mathbf{t}) = (I-\alpha A)^{-1}\mathbf{t} = \sum_{k=0}^{+\infty} \alpha^k A^k \mathbf{t} \tag{7}$$

In (7), the matrix $A = D_t D_p^{-1} G$ (see Section 3) where D_t and D_p are diagonal matrices of individual productivities and the number of projects involved in during the period under consideration. The construction of network G is as described above. We proxy individual productivity by the quantity and quality of their overall output. We use a measure of Euclidean index proposed by Perry and Reny (2016). For each i, the productivity $t_i(\tau)$ at period τ is the Euclidean index of the history of i's publications until year τ, normalized by seniority. Let $P_i(\tau)$ denote the set of all i's publications until τ, and let c_p be the total citation count for a typical $p \in P_i(\tau)$. Then

$$t_i(\tau) = \frac{1}{\tau}\left(\sum_{p\in P_i(\tau)} c_p^2\right)^{\frac{1}{2}} \tag{8}$$

An author's Euclidean Index hence monotonically increases in the number of publications, given that each publication is cited at least once.

The alternative measures would be the total number of publications or the total citation count normalized by the number of years of experience. The Euclidean index however takes into account both measures making it a more accurate measure. To compute each $t_i(\tau)$, we use data obtained from Scopus for about 85% of the nodes of each of the social networks of intellectual collaboration.

The number of projects each researcher is involved in each year is the number of publications in the current year and the next year. To account for shared work, each project is divided by the number of co-authors. For example for someone that published one single authored article in 2005, and one co-authored article in 2006, the number of projects in 2003 would be 0, in 2004 1, in 2005 1.5, and in 2006 0.5.

The parameter α is generally referred in the literature as *attenuation factor*. It discounts the distance between agents that are not collaborating directly. The smaller α, the less impact

[23] Without loss of generality, we take $\beta = 1$ such that $\alpha_\beta = \alpha$, and $\mathbf{e}^* = \mathbf{b}(A,\alpha,\mathbf{t})$.

distant agents have on an agent's equilibrium effort. In our model, α captures the contribution of complementarities in collaboration to ones research output. Proposition 1 provides equilibrium conditions; that is $\alpha < \frac{p_1}{t_n \mu_1(G)}$, where $\mu_1(G)$ is the leading eigenvalue of the weighted network G. This condition can however be equivalently stated as $\alpha < \frac{1}{\mu_1(A)}$ where $\mu_1(A)$ is the leading eigenvalue of the weighted network A. In the regression, we use $\alpha = \frac{0.99}{\mu_1(A)}$, and for robustness we also use $\alpha = \frac{0.95}{\mu_1(A)}$ and $\alpha = \frac{0.90}{\mu_1(A)}$.

FIGURE 5 ABOUT HERE

FIGURE 6 ABOUT HERE

In real world networks it is common that nodes are not connected, not even indirectly. Formally, two nodes belong to the same network component $\updownarrow$ if there exists an alternating sequence of nodes and ties, called a path, between them. There can be as many components as there are nodes if all nodes are isolated. The size of a component is the number of nodes (i.e. the number of academics) it contains. The component containing the most nodes is called the giant component.

Figure 5 shows the social network of informal intellectual collaboration in 2011. We contrast it with a pure co-author network for the same period as shown in 6. While the co-author network is unconnected and split into many different components, the network of intellectual collaboration is much larger and consists of one big component only.[24]

For the regression, we compute and use the Bonacich centrality for all nodes in the giant component of A and omit the other components for technical reasons. This is mainly centralities across components are not comparable: If a node i belongs to a small network component, all other nodes are fairly close. In contrast, a node in a large network might have a potentially much smaller centrality because many other nodes are far away.

TABLE 7 ABOUT HERE

Table 7 shows the evolution of the social network of informal intellectual collaboration over time. To examine non-overlapping networks only, we list the networks of informal intellectual collaboration for the years 2002, 2005, 2007, 2009 and 2011. Between 2002 and 2011, the network has nearly doubled, from 3557 nodes to 6997 nodes. The largest component, the giant component, captures a slightly greater share of the network today, increasing from 3377/3557 = 94.9% in the 2002 network to 6750/6997 = 96.4% in the 2011 network. The error from using the giant component only to compute centralities is hence relatively small.

[24]Typically, high Bonacich central nodes are clustered together meaning that Bonacich centrality points to the best connected nodes in a network. That is, it is highest when an academic is connected to academics with high Bonacich centrality themselves. We explore the determinants of centrality in the social network of intellectual collaboration in greater detail in Georg and Rose (2015).

In general, the network today is less connected: Though there are more edges connecting the nodes, there are more distinct components, starting with 37 in 2002 and arriving at 54 today. Network density, the ratio of the total number of links to the total number of all possible links, has decreased from 0.002 to 0.0014 and the diameter has increased by 3 nodes (from 14 to 17). The diameter is the longest of all shortest paths in the giant component. It can be understood as maximum distance from one side of the network to the other.

6 Identifying the Impact of Informal Collaboration on Publication Success

Following our structural model of section 2 we hypothesize that the scientific impact of a research article in financial economics–i.e. how well it is published and how often it is cited–increases in the Bonacich centrality of its commenters. Bonacich centrality (equation (7)) is measured in a social network of informal intellectual collaboration where links are weighted by the product of frequency of the exchange and productivity of the researchers involved. Productivity is the Euclidean Index of a researcher's publications ((8)) divided by her experience, which is the number of years since the first publication.

6.1 Identification Strategy and Model Setup

The key identification challenge we face when studying whether input from colleagues is more beneficial if the colleagues are more central is that social networks are inherently endogenous (Manski, 1993; Graham, 2015). That is, one may observe a change in an agent's outcome (e.g. how many citations one of her research articles receive three years after publication) due to the outcomes of the agent's network neighbors. The social tie between the two agents is, however, subject to decisions taken by the agents. Specifically, there might be a social tie between two researchers because they are similar, which results in similar outcomes. Or the tie might be the result, not the cause, of individual efforts. Manski (1993) labels this problem the reflection problem.

We address this problem through several measures. To control for author ability, we add citation information from Scopus. Scopus provides a large database with individual metrics, which we use to compute the number of yearly publications of an academic, as well as her yearly number of citations. We use the first recorded publication to estimate experience, measured as the difference between the year in which the first publication was published and the year of publication of the research article under consideration. To capture non-monotonic age-effects (Levin and Stephan, 1991), we include the square of the author seniority. For groups of authors we summed the experience of all authors. To control for author ability, we use Euclidean index of an author's publications until a given year, as defined in (8).[25]

[25] As an alternative measure we use the simple citation stock, resulting only in lower p-value of the main variable of interest.

Our estimation uses two samples. In the *NBER sample*, we restrict ourselves to research articles that a) have been presented at finance-related NBER summer institutes since 2001,[26] and b) were published in one of the three top Finance journals The Journal of Finance, The Review of Financial Studies, and the Journal of Financial Economics. We look at centralities of discussants. Finally, in the *commenter sample*, we look at research articles from the *full sample* with at least one commenter. When we use this sample, we consider the centralities of all commenters.

The hypothesis in all samples is the same: The number of citations to a research article increases in the connectedness of its informal collaborators. To test this hypothesis, we estimate the following model:

$$\begin{aligned} Citations = \alpha_0 + \alpha_1 \cdot ArticleCharacteristics + \alpha_{\mathbf{2}}\mathbf{D}_{journal} + \alpha_{\mathbf{3}}\mathbf{D}_{year} \\ + \beta_0 \cdot AuthorCharacteristics + \beta_1 \cdot DiscussantCharacteristics \\ + \beta_2 \cdot Centrality \end{aligned} \tag{9}$$

where $ArticleCharacteristics$ is a vector of article-specific variables that contain the number of authors and the number of pages. $\mathbf{D}_{journal}$ is a vector of dummy variables that captures journal fixed effects such as editor skills and preferences or management policies, while $\mathbf{D}_{year}$ represents publication year fixed effects. $AuthorCharacteristics$ and $DiscussantCharacteristics$ are vectors controlling for author and discussant characteristics and include Euclidean Index, the number of projects, the experience and the squared experience of authors and discussants, respectively. In case multiple authors have co-authored or multiple discussants have discussed the research article, we take the total of each author's or discussant's values. Since we also control for the number of authors and the number of discussants, each of the characteristics variables effectively captures the mean effect. Author characteristics were counted in the year before publication while for discussants the characteristics were counted in the year of the discussion.

Our main contribution is to introduce $Centrality$, the rank according to Bonacich centrality of discussants in the social network of intellectual collaboration in finance, as define in (7). Following our model of scholarly collaboration, the Bonacich centrality of a discussant is proportional to her effort. We use ranks rather than values because absolute values are meaningless. It is not valid to infer qualitative comparisons from the values rather than on an ordinal scale. Again, if more than one discussant discussed the same manuscript (a rare event) we take the total while we control for the number of discussants.

The relevant social network is always inferred from the acknowledgments section of all finance research articles published in the three years preceding the article publication, excluding the year of publication itself. This has the advantage that the relevant social network is constructed without the article under consideration. Hence the author's own network status does not influence the positions of the commenters.[27]

[26]We include the summer institutes of the following NBER groups: Asset Marketing/Real Estate (AMRE), Asset Pricing (AP), Corporate Finance (CF), Capital Markets in the Economy (EFEL), Household Finance (HF), International Finance & Macroeconomics (IFM), Monetary Economics (ME), and Risk of Financial Institutions (RISK)

[27]A possible exception is when authors and commenters have collaborated in previous years, which is not very

To assure comparability across networks, we compute the Bonacich centrality with an attenuation factor of $0.99 \times 1/\mu_1(A_t)$, where $\mu_1(A_t)$ is the leading eigenvalue of the adjacency matrix of network A for year t. The attenuation factor governs how quickly externalities from actors to a focal decrease in their distance to the focal node. By using a constant scaling of the inverse of the leading eigenvalue our Bonacich centrality measure takes into account the dynamics of the networks. Using different scaling factors do not change the results, as we show in sensitivity analysis in section 7.1.

Missing Bonacich centrality ranks are mean-imputed on the article level, i.e. only for articles where Bonacich centrality ranks are missing for all of the discussants. A rank can be missing because the discussant is not part of the network (i.e. she is not authoring an article or not acknowledged as ordinary commenter in any of the six journals during the last three years), or because she is not part of the largest component. By definition, Bonacich centralities rely on a connected component and only hold for this component. Though it is technically possible to compute centralities for each node in each component, centralities of different components are not comparable.

Our main dependent variable is the count of citations since publication, which we obtain from Scopus for all research articles in our sample. The distribution of citations is skewed, discrete and non-negative. Therefore we estimate a negative binomial regression model (Mullahy, 1986). Being a generalized linear model, the parameters are evaluated at sample mean, i.e. holding all variables fixed at their mean values. To ease interpretation, we compute and present marginal effects, i.e. the coefficients we present are the expected *percentage* increase in the outcome variable when the explanatory variable increases by 1 unit and when all other variables are held constant at their mean and all dummy variables at 0.

6.2 A Quasi-Natural Experiment: Discussants at NBER Summer Institutes

We identify a network effect exploiting a quasi-natural experiment that occurs frequently in academic life and which addresses the reflection problem: The assignment of discussants at conferences. We focus on a select group of research articles, namely those presented at NBER summer institutes. Our main explanatory variable is the eigenvector centrality rank of discussants in the social network of intellectual collaboration three years prior to publication.

The US-based National Bureau of Economic Research organizes its work through a number of working groups that meet on a yearly basis. These meetings (called summer institutes) are small workshops with an extremely low acceptance rate. Usually less than 10% of the submitted research articles are accepted for presentation. Nearly all of the presented research articles are discussed by an external discussant, i.e. a discussant that is not an author at the same meeting. These discussants are usually hand-picked by the group coordinator and very rarely reject an offer to discuss a research article. Discussants are hence assigned exogenously from the author's perspective which allow us to addresses the reflection problem.

---

common for discussants.

[TABLE 8 ABOUT HERE]

The list of every group meeting that has been held since 2001 is available online, including the title of each research article, the names of the authors, presenters and discussants. We focus on finance-related working groups that have existed between 2001 and 2011. Specifically, we consider the summer institutes for the Monetary Economics (ME), International Finance & Macroeconomics (IFM), Corporate Finance (CF), Asset Pricing (AP), Capital Markets in the Economy (EFEL), Risk of Financial Institutions (RISK), Household Finance (HF), Finance & Macro Meeting (MEFM), and Asset Marketing/Real Estate (AMRE) groups. Between 2001 and 2011, a total of 550 presentations took place at a total of 44 summer institutes, including double counts (i.e. manuscripts presented twice). However, not every presentation eventually resulted in a publication. Overall, 417 (76%) of the presentations resulted in publication in 59 different journals or books until January 2017.[28] Table 8 gives an overview of the number of presentations by year and NBER group. The ratio of articles published over the number of manuscripts presented is highest for NBER groups Asset Pricing and Corporate Finance (both 83%) and lowest for Household Finance (54%). The most important outlets for the published research articles are The Journal of Finance (65 research articles), followed by The Review of Financial Studies (51), the Journal of Financial Economics (45) and The American Economic Review (34). 10 articles were published in books.

Of the 417 presented manuscripts that were eventually published, 27 were presented multiple times, so that the final number of observations in our sample is 389. For all the 389 research articles, we obtain the number of total citations up until April 2017 as well as bibliographic measures from Scopus. The dependent variable is the count of citations since publication. In addition to the variables outlined in 9, we control for *age*, the number of years between the first NBER summer institute a research article was presented in and the publication date, and its square, age^2.

We present results in three different samples. Our main sample consists of 161 articles that were published in a top Finance journal, i.e. The Journal of Finance, the Journal of Financial Economics, and The Review of Financial Studies. The reason for this is that we get a homogeneous set of articles, i.e. articles primarily intended as general interest finance articles rather than e.g. as policy articles. We control for these journal-specific effects such as editorial policies or journal idiosyncrasies (e.g. the JFE offers a fast-track option for articles) with fixed effects relative to The Journal of Finance. The second set named *NBER(+Econ)* extends the set of journals we look at by the top 5 Economics journals, The American Economic Review, Econometrica, the Journal of Political Economy, the Review of Economic Studies, and the Quarterly Journal of Economics. 344 articles published in journals that occur at least three times in the sample. Finally the third set includes all publications regardless of the journal or book they appeared in. For example, a total of 10 articles were published in books and further 42 articles were published in journals with less than three publications in our sample (for example, out of the 389 presented manuscripts of our sample, only one was published in Brookings Papers of

[28] Some research articles changed the title. We, therefore, conducted an internet search for each article based on the authors and abstracts to identify those research articles with a changed title.

Economic Activity). In each specification we control for these with fixed effects relative to The Journal of Finance.

The full set of 389 articles have been written by 629 distinct author and discussed by 261 distinct discussants. For each of them we compute the Euclidean Index, experience and the number of projects using data from Scopus. We estimate the experience of an author or a discussant as the number of years since first publication. We estimate the number of projects as the number of publications in the current and the next year, each divided by the number of authors.

[TABLE 9 ABOUT HERE]

Table 9 presents the 25 most Bonacich central discussants of published manuscripts for four different periods. Each column refers to one social network of informal intellectual collaboration. The first column refers to the years 2000-2002, the second to 2003-2005, the third to 2006-2008 and the last to 2009-2011. Numbers in parenthesis indicate the number of times they have discussed a research article in the *NBER sample* (i.e. an article that was eventually published until March 2017).

[TABLE 10 ABOUT HERE]

Table 10 presents all summary statistics for all continuous variables for the NBER sample. The average research article in this sample has garnered 85 citations since publication, has been written by 2.4 authors, consists of 34.8 pages, was discussed once and has been published 3.3 years after the first presentation at an NBER summer institute. The authors of the average research article have a joint Euclidean Index of 485.6 citations as measured in the year before publication. They were were engaged in 21.3 projects and have a joint experience of 23.8 years. On average, the sum of Euclidean Indices of the discussants of a research article's amounts to 187.3 in the year of the discussion, which is considerably lower than the authors'. The discussants were engaged in 2.1 projects. The joint experience of all discussants amounts to an average of 12.1 years prior to the discussion, ranging from -5 to 35 years (a negative number implies that the discussant published her first article after her discussion). The vast majority of articles has been discussed once, which partly explains the huge differences in the totals for authors and discussants.

[TABLE 11 ABOUT HERE]

Table 11 reports Spearman and Pearson correlations between all variables used in the NBER sample. We use them to test for two things. First, we want to check whether there is assortative matching between authors and discussants, i.e. if articles by more senior and productive authors are discussed by more senior and productive discussants. And second, whether discussant network characteristics are highly correlated with how senior and productive the discussant is.

On the first count, we find low correlation between authors' Euclidean Index and total experience and the characteristics of discussants, which indicates that there is no assortative mixing. On the second count, we find weak correlation between discussant eigenvector centrality rank and other discussant characteristics (i.e. their seniority and citation stock). This indicates that centralities of discussants could capture information not captured in traditional measures of commenter quality. Interestingly, we find a negative correlation between discussants' total experience and their total Bonacich centrality rank (Pearson: 0.26, Spearman: 0.20), i.e. more experienced discussants are more likely to be *less* Bonacich central.

[TABLE 12 ABOUT HERE]

We explore the relationship between author metrics and discussant characteristics in a separate negative binomial regression to further check for assortative matching between authors and discussants, shown in Table 12. The unit of observation is the discussant-presentation combination, i.e. a discussion by two discussants results in two observations, such when the same author discusses two manuscripts (The values for N differ from 389 when the dependent variable was not available for this discussant.). We run separate regressions for the three different discussant characteristics we use: the discussants' joint Euclidean Index, the joint number of projects, the joint seniority, and the cumulated Bonacich centrality rank. In the sensitivity analysis we test whether the betweenness centrality of discussants is associated with a higher citation count, which is why we include a discussant's betweenness centrality rank here. Each variable is measured in the year of the discussion. For the first three variables we expect positive signs while for the two centrality measures we expect a negative sign, because they are ranks.

For discussants' Euclidean Index, no coefficient displays statistical significance. This indicates that manuscripts authored by researchers that have a higher Euclidean index themselves (i.e. are engaged in more projects, are more experienced or simply are more numerous in terms of coauthors) are not matched with more productive discussants. The same holds for a discussant's number of projects. Regarding discussant's experience we see a statistically weak, but negative correlation between author group size and more experienced discussants: Each author more from the sample mean results in a discussant that is 9.3% (i.e. $0.093 \times 11.79 \approx 1.1$ years) less experienced. Regarding position in the network, there is some statistical evidence that author groups with a higher joint Euclidean Index receive *less central* discussants, as the respective coefficients are positive. The statistical correlation is stronger for betweenness centrality ranks of the discussant.

By contrast, there are statistically significant relationships between some of the NBER groups and some discussant characteristics.[29] We take this as evidence that more manuscripts with more productive authors do not automatically get better discussants, where "better" is understood as more senior, more productive, and better positioned in the network.

[29] Detailed results are available from the authors upon request.

[TABLE 13 ABOUT HERE]

Table 13 reports the results of our main specification (9). In all regressions, we control for the set of article characteristics (the number of authors and the number of pages), which we do not report due to space restrictions. We include fixed effects for the journal and the publication year.[30] Column (1) serves as reference model and does not include any discussant characteristics or discussant Bonacich centrality ranks. The combined Euclidean Index the authors has no economically or statistically significant effect on the total number of citations an article receives, which is an interesting finding in itself. One explanation is the relatively high share of manuscripts written by young authors. In contrast, the authors' combined experience is both statistically and economically significant: A one year increase from the mean is associated with 3.5 (~ 3) more total citations at the sample mean. The effect is non-monotonic. This can be seen from the statistically significant negative coefficient for squared experience. Column (2) introduces discussant characteristics such as discussants' combined Euclidean index. None of the coefficients are statistically significant.

Column (3) of table 13 introduces discussant Bonacich centrality rank. The coefficient is statistically significant with a p value equal to 0.016. Each rank increase from the mean is associated with an increase of 0.3% (~ 0.026) citations at the sample mean. Given that the networks are very large and there are about 6000 ranks in each of the networks and that the sample mean of discussant Bonacich rank is at 412.2, the effect is economically significant. For example, replacing a discussant with Bonacich centrality rank of 450 with one having rank 350 is associated with $0.0003 \times 100 \times 85 \approx 2.55$ more citations, all else equal. At sample mean this corresponds to a 3% improvement in the number of citations.

Column (4) and (5) are alterations of the regression. Column (4) augments the sample with articles that were published in five general interest journals, with the average total citations being equal to 84.5 citations. Column (5) extends the sample to all journals, where the sample mean of total citations equals 67.4. When looking at top finance and top economics journals (column (3)) the effect of discussants' total Bonacich centrality decreases by about a third to -0.0002. This corresponds to a citation increase of 1.7 citations at sample mean for a Bonacich centrality improvement by 100 ranks. Once we include all journals, the coefficient decreases to -0.0001. In this sample a 100 rank improvement in the discussants' joint Bonacich centrality for the average article corresponds to 0.6 more citations. This implies that the effect of discussant centrality is larger for top journals than for non-top journals.

Our model predicts that research impact increases in the effort put forward by informal collaborators. The empirical results support this relationship. We find positive and statistically significant relationship between Bonacich centrality and research impact measured as citation count.

Another measure of impact is the publication in a top journal. We asses the impact of discussant centrality on publishing in a top journal using a logit regression where the dependent variable equals 1 if the article was published in a top journal, and 0 otherwise. Top journal

[30] The extended regression tables are available from the authors upon request.

refers to the set of three top finance journals: The Journal of Finance, The Review of Financial Studies, Journal of Financial Economics.[31] For this estimation we use the *NBER sample* with 389 observations.

[TABLE 14 ABOUT HERE]

[TABLE 15 ABOUT HERE]

Table 14 reports summary statistics for the extended *NBER sample*, while Table 15 reports Spearman and Pearson correlation coefficients. 40% of the articles in this sample were published in a top Finance journal. Mean values in this sample largely correspond to mean values of the sample used for the Negative Binomial regression, i.e. that which consists of top finance publications only. The only exception is that here the authors are less productive, as seen from the Euclidean Index which is about 15%. The empirical probability of publishing in a top journal is not strongly correlated with any of the article characteristics, author controls or discussant controls. None of the absolute Spearman (Pearson) correlations exceeds 0.20 (0.22) which is the correlation with the number of pages. Interestingly, the number of discussants, age, author experience, discussants' total projects and discussants' total experience are negatively correlated with top publication probability, albeit very weakly. However discussants' total Bonacich centrality rank and probability of publishing in a top journal correlated stronger. The respective Spearman correlation coefficient equals -0.34 (Pearson: -0.35).

[TABLE 16 ABOUT HERE]

Table 16 reports marginal effects for this regression. A marginal effect is the expected percentage increase of the dependent variable from the mean if the independent variable increases by one unit and all other variables are fixed at their respective means.

Column (1) of table 16 is the baseline case with author characteristics and article controls. All author controls are highly statistically significant, except for the authors' joint number of projects. However, the coefficients are negative, implying a negative relationship between a higher joint citation stock and top journal publication probability. The same is true for joint experience, where there appears to be a non-monotonic relationship. In column (2) we add discussant controls. As previously in the regression for total citations, none of the discussant characteristics display statistical significance. Finally in column (3) we add discussants' total Bonacich centrality ranks. We expect a negative sign as a higher rank would indicate a less central position. The effect is indeed negative and highly statistically significant: For each rank increase the likelihood of publishing in a top journal increases by 0.1%, all else fixed at sample mean. Replacing a discussant with rank 1000 by a discussant with rank 900 increases top publication likelihood by 10% at sample mean–a huge difference in the highly contested space for publications in top journals.

[31] Again, the inclusion of the top five Economics journals The American Economic Review, Quarterly Journal of Economics, Journal of Political Economy, Econometrica or Review of Economic Studies as a robustness check does not change our results.

6.3 A Comparison With the Full Sample of Commenters

We contrast our analysis above with an analysis of the full *commenter sample*, i.e. the unrestricted sample including all commenters from publications in a set of six major finance journals published between 1998 and 2011 that have acknowledged at least one commenter (N=4406). These journals are The Journal of Finance, The Review of Financial Studies, the Journal of Financial Economics, the Journal of Financial Intermediation, the Journal of Money, Credit, & Banking, and the Journal of Banking and Finance. In this sample, we include all commenters acknowledged on a research article (except discussants and referees). In addition to the variables used in equation 9, we control for *topic* and other measures of *informal intellectual collaboration*. We proxy the topic using a matrix of binary dummy variables for all one-digit JEL codes of the general category G (financial economics) to filter topical effects (i.e. G0, G1, G2 and G3). Measures of informal intellectual collaboration include the number of commenters, seminars and conferences acknowledged, as well as whether the research article has been presented at a top conference. To each article we assign the total of the Bonacich centrality ranks of all acknowledged commenters. Bonacich centrality is computed according to (7), i.e. as in the structural model. We control for the sum of the commenters' Euclidean Index of citations in the year before publication, and their joint experience measured in number of years since first publication. We obtain the information on citations from Scopus with a coverage of at least 84.02% percent of all nodes in the network's giant component (because only the giant component is relevant for the centrality computation). The dependent variable is again the count of citations until January 2017 and obtained from Scopus in the same month.[32]

[TABLE 17 ABOUT HERE]

Table 17 presents summary statistics for the commenter sample. In this sample, the average research article has garnered 73.8 citations until January 2017, has been written by 2.2 authors and consists of 26.7 pages. The average research article additionally acknowledges 9.0 commenters, 5.2 seminars and 2.8 conferences. 20% of the articles in this sample have been presented at a top conference. The authors' combined Euclidean Index equals 214.5, which is about half the figure of the *NBER sample*. The authors were engaged in a total of 3.7 projects and have a joint experience of 20.7 years. Negative experience in author experience are due to the construction of the variables, which are measured in the year before publication. If for example an article is the first for all authors, they have a negative experience. This is the case for 15 publications.

[TABLE 18 ABOUT HERE]

Table 18 reports Spearman and Pearson correlations between all variables used in the commenter sample. Pearson correlation coefficients are depicted along the lower triangle and

[32] Using citations according to Web of Science, which are usually lower, does not change the results.

Spearman correlations along the upper triangle. As in the previous sample, there is only a positive weak relationship between an article's citation count and author ability measured as sum of Euclidean indices. The respective Spearman correlation coefficient is 0.16 (Pearson: −0.21). Correlation coefficients of total citations with the number of authors or the number of pages are equally weak (Pearson: 0.07 and 0.24, respectively. Spearman: 0.46 and 0.37, respectively). We find a positive weak correlation between measures of informal collaboration (number of commenters, seminars or conferences) and measures of article quality (total citation and top publication probability). The highest Pearson correlation coefficient is 0.32 and between top publication probability and number of commenters (it is the same for the number of seminars). The highest Spearman correlation coefficient is 1 between top publication probability and the number of seminars. Spearman correlation furthermore indicates no correlation between top publication probability and number of conferences. Finally, table 18 shows again a weak relationship between measures of commenter connectivity and any other variable, as Pearson coefficients do not surpass −0.23 (note that a negative coefficient indicates a positive relationship because the total Bonacich rank is ordinal). Surprisingly, Pearson correlation indicates a negative relationship between commenters' total Bonacich rank and top publication probability. However, coefficients indicate a stronger correlation between commenters' total Bonacich centrality rank and their combined experience (Pearson: 0.42, Spearman: 0.19).

There is a very weak correlation between authors' total citations and all measures of informal intellectual collaboration, but a very weak negative correlation between authors' total experience and the number of commenters (−0.09). More senior authors hence tend to ask fewer colleagues for input on a manuscript.

TABLE 19 ABOUT HERE

Table 19 reports marginal effects for all variables. We control for journal, publication year and topic in each of the specifications. Column (1) serves as reference model and excludes any commenter-related variables. All variables except the number of conferences and the authors' combined experience are statistically and economically significant. This includes the number of acknowledged commenters and the number of acknowledged seminars. For example, going from the sample mean of 9 commenters to 10 commenters is associated with a 1.6% (~ 1.16) increase in the number of citations, holding all other variables fixed at their mean. Each additional seminar from the mean is associated with 1.5% (~ 1) more citations.

In column (2) of Table 19 we add the commenter characteristics at an aggregated level, i.e. their total Euclidean index, their total number of ongoing research projects, and their total experience as measured as number of years since the first publication, and the square thereof. All coefficients are statistically highly significant. That is, acknowledging more productive commenters is associated with a higher citation count, and so is acknowledging commenters which generate spillovers from their ongoing research projects. However, the coefficient for combined experience is negative, indicating that more experienced commenters are associated with a lower citation count, but at a decreasing rate.

Column (3) of Table 19 finally adds the sum of Bonacich centrality ranks of all acknowledged commenters to estimate the effort brought forward to the research project. The coefficient is statistically highly and significant and has a value equal to 0.0001. The mean of this variable is very high with 13192.4. Reducing the combined Bonacich centrality rank for the average publication by 100 ranks is associated with a citation increase of 0.01% (0.725) citations at sample mean. Other coefficients remain unchanged.

Finally, we again assess the importance of informal intellectual collaboration with central colleagues for publishing in top journals. We then can estimate the probability of top journal publication. Table 20 presents the results of a logit regression as marginal effects, showing the percentage increase in the dependent variable when the explanatory variable is increased by 1 unit and all the other variables are held constant at their mean. We again control for publication year and topic via the listed JEL codes.

TABLE 20 ABOUT HERE

Column (1) of Table 20 serves as reference model and excludes any commenter-related variable. All variables except author total experience and the square thereof are statistically highly significant. Interestingly, the coefficient for the number of conferences is negative, indicating that more conferences are associated with a lower acceptance probability at top journals. This might be an indication for the fact that publications in field journals tend to be presented at a larger range of more specialized conferences, while publications in the top finance journals tend to be presented at the relatively smaller number of top conferences in financial economics. The variables with the highest effects are the number of seminars (19.4%), the number of authors (35.6%), and top conference presentation (143.3%). Presenting at a top conference is associated with doubling the likelihood of publishing in a top journal. Each additional commenter is associated with a 6.6% higher acceptance probability at sample mean. The lack of statistical significance for author experience indicates that young authors have neither a malus nor a benefit when it comes to publishing in top finance journals.

Column (2) of Table 20 introduces commenter characteristics at an aggregated level: Their total Euclidean Index, their total experience as measured as number of years since the first publication, and the square thereof. Other coefficients change little, except for the number of commenters, whose coefficient doubles in size. The coefficient for the total Euclidean Index is statistically highly significant, indicating that informal collaboration with more productive researchers increase top journal publication probability. Unlike in the assessment for the count of an article's citation count, the coefficient for the count of ongoing projects is negative. This indicates negative spillovers when article impact is measured in terms of top journal publication probability. Again, the coefficient for combined commenter experience is negative but statistically insignificant.

In column (3) of Table 20 we add the total of all commenters Bonacich centrality rank. The coefficient is highly statistically significant with a p-value equal to 0 and has the expected negative sign: An increase from the mean by 100 ranks of all commenters is associated with a 1% increase in top journal publication probability. This is significantly lower than the corresponding increase in the NBER sample and is consistent with our interpretation that indeed complementarities and information spillovers lead to a higher scientific impact: a discussant is usually someone who is an expert in the area of the paper who spends a substantial amount of time on her discussion. It is likely that there is more information spillover than in a more casual meeting where one researcher provides commentary to the other.

Even though this sample reports statistical significant correlations rather than the direction of the effect (For example, it might as well be that authors that publish in top journals collaborate more informally), the findings are consistent with existing literature. For example Brown (2005) compares accepted and rejected manuscripts to The Accounting Review and finds that informal collaboration of all kind –seminar presentation, conference presentation, commenters– increases acceptance probability. We explore the connection to existing studies in the field in greater detail in section 7.3.

7 Discussion and Conclusion

This section performs a sensitivity analysis and compares our sample with existing studies of informal intellectual collaboration.

7.1 Alternatives to Bonacich centrality

Our main explanatory variable is the rank according to the Bonacich centrality, summed over all discussants. Bonacich centrality is defined as (7). To test the robustness of our results, we alter the attenuation factor of the Bonacich centrality. Table 21 presents the results. In each column we test the main specification used in column (2) of Table 13, but replace the Bonacich centrality with alternative values of the attenuation factor.

[TABLE 21 ABOUT HERE]

In our main specification, with results in Table 13, the attenuation factor equals $0.99 \times 1/\mu_1(A)$, where $\mu_1(A)$ is the leading Eigenvalue of the network A under consideration. In Table 21, column (1) the attenuation factor equals $0.95 * 1/\mu_1(A)$, while in column (2) the equals $0.90 \times 1/\mu_1(A)$. Compared to the main specification, the coefficient remains the same at 0.0003. Hence our results are robust to changes in the attenuation factor.

In column (3) we compute the Bonacich centrality without any link weights. The coefficient changes signs and looses statistical significance. We take this as sign that accounting for the frequency of interactions among scientists indeed matters in capturing knowledge spillover.

In column (4) we compute the standard Eigenvector centrality, which is a special case of the Bonacich centrality. When the starting vector $\mathbf{t}$ is set to 0, so that there are no initial centralities, and the attenuation factor is set equal to $1/\mu_1(A)$, an equivalence exists between

Bonacich and Eigenvector centralities. The results are virtually unchanged, with a highly statistically significant coefficient of -0.0003. We interpret this result as implying that the significance of weighted Bonacich centrality $\mathbf{b}(A, \alpha, \mathbf{t})$ in our main results is not influenced entirely by the vector of types $\mathbf{t}$, but rather purely as a result of the collaboration network.

The above results show that connections in the network of informal intellectual collaboration matter. In column (5) we change the network definition such that co-author links are included. While previously researchers were only connected when one acknowledges the other, in the combined network researchers are connected when one acknowledges the other and/or whenever they jointly publish an article (see Section 5.4 for details). Again, the coefficient is virtually the same as in our main specification. In our model, and in particular the equation for output (1), we do not differentiate knowledge spillover among scientists based on whether the link is formal or informal collaboration. We assumed that both matter. The results for a combined network seems to suggest that such a distinction may not be necessary, implying that the network of informal collaboration sufficiently captures the structure of intellectual collaboration in research

7.2 Betweenness centrality

The equilibrium behavior of our model shows that Bonacich centrality best captures individual influence in an environment of intellectual collaboration with positive complementarities. The existence of complementarities is necessary for Bonacich centrality to be an appropriate measure of influence in equilibrium. If the underlying process driving our empirical results were pure information flow, then other centrality measures can equally capture a scientist's level of influence. In particular, the *between centrality* is so often used to measure the level of individual influence in relation to information flows within a network.

Betweenness centrality was introduced by Freeman (1978), and is defined as the probability that a node is on a shortest path between any two nodes in the network. Formally, let i, j and $k \in N$ be academics in a connected component l (i.e. there exists a path between all nodes). Denote the shortest path between j and k as σ_{jk} and the number of shortest paths between j and k that contain node i as $\sigma_{jk}(i)$. Betweenness centrality $C_B(i)$ of node i is then given as:

$$C_B(i) = \sum_{j,k \in N} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \tag{10}$$

Unlike Bonacich centrality, the betweenness centrality score has a direct interpretation: It is the number of shortest paths that go through a node divided by the number of all possible shortest paths. A higher centrality score then implies that the node is part of more shortest paths. For the process of information flows within the network, agents with the highest betweenness centrality receive more and a variety of information than those with lower centrality.

Although both measures captures an agent's influence in terms of information possessed in the process of information flow, Bonacich centrality is unique to the process of strategic interactions with positive complementarities. Hence, if scientists' betweenness centralities significantly influence their research output, then our results above could also be a result of pure information contagion and not existence of positive complementarities.

[TABLE 22 ABOUT HERE]

In Table 21 we test whether betweenness centrality of discussants is statistically significantly associated with an article's citation count. In each column we test the main specification used in column (2) of Table 13, but do not use the Bonacich centrality derived from the model.

Column (1) uses all discussants' joint betweenness score and column (2) uses all discussants' joint betweenness rank as explanatory variables. We have normalized the betweenness score as defined in (10) at the article level for convenience, such that a unit increase corresponds to an increase by one standard deviation. The coefficient for betweenness centrality score is not statistically significant but has the expected sign. This indicates that the positive and significant result we obtain for Bonacich centrality is at the very least partially driven by strategic complementarities. Note that, while centrality rank is the appropriate measure for Bonacich centrality, due to the lack of a direct interpretation of the centrality values, this is not true for betweenness centrality. Here, the values themselves are the most appropriate unit of observation.

7.3 Relation to Existing Studies of Informal Intellectual Collaboration

Laband and Tollison (2000) and Brown (2005) estimate the impact of informal intellectual collaboration on the number of citations of published research articles. Laband and Tollison (2000) use 251 featured articles published in the Review of Economics and Statistics during the years 1976-1980. They estimate the effect of the number of acknowledged commenters to explain the number of citations in the following six years to that article. They control for the cumulative stock of citations from the previous five years for all authors, as well as the number of pages. They show that number of commenters is statistically significant and positive. In alterations to the model they add the caliber of commenters (i.e. their joint citation stock) and how many of the commenters are to colleagues from the same department or from away.

Brown (2005) uses a negative binomial regression similar to ours and a sample of 256 research articles published in The Accounting Review, the Journal of Accounting Research, and the Journal of Accounting and Economics during 2000-2002. The dependent variable to measure publication success is the number of citations since publication according to the Social Science Citation Index. His main explanatory variables are the number of commenters, the number of conferences, and the number of seminars. Brown (2005) controls for the number of pages, the number of authors, whether the research article was highly downloaded from SSRN, and also uses journal- and time- fixed-effects. He finds that only seminars have a statistically significant and positive impact on publication success. Estimating the impact of acceptance probability on the journal he edited – The Accounting Review – he finds that all forms of informal intellectual collaboration matter.

We aim to replicate both studies' results, with two slightly different variable definitions: First, we use citations over the entire lifespan of an article (instead of six years) as dependent variable. Second, we measure author quality as the total citations, not only in previous five years. Moreover, due to lack of data, we do not control for SSRN downloads when we aim to replicate Brown (2005). We estimate both regressions on our *commenter sample* where we set the number of acknowledged commenters (seminars, conferences) to 0 if none are acknowledged. Results of this estimation are shown in Table 23.

[TABLE 23 ABOUT HERE]

Columns (1) through (3) in Table 23 replicate model (1) through (3) of Table 4 of Laband and Tollison (2000), while column (4) replicates Table 8C of Brown (2005). Unlike Brown we find a statistically significant relationship between the number of authors and citation count, as well as between the number commenters and citation count, even after controlling for the number of acknowledged seminars and conferences. Moreover, we find a statistically significant negative relationship between article length and citation count, where Brown finds a statistically significant positive relationship. However, the main result, namely that informal collaboration increases subsequent citation count, holds. We take this as confirmation that our data are akin to the data used in the existing literature.

7.4 Conclusion

Researchers that collaborate either formally or informally inevitably diffuse information for example on new ideas, emerging trends and upcoming challenges. The extent to which individuals participate in the diffusion process depends on their position in the social network of intellectual collaboration (Jackson, 2014). We show that research articles benefit more from a discussant's comments when the discussant is more Bonacich central, i.e. when she is closer to the most connected clique in the network (Bonacich, 1987). An increase by 100 ranks in eigenvector centrality from the mean is associated with an increase by ~ 2.55 citations for the average article. Given that the average article had a discussant ranked at position 412 and that the network consists more than 5000 researchers, our results suggest a high scope for improved information access.

We find a statistically significant correlation evidence for network effects in the probability of publishing in a top journal if we consider the sample of all commenters, although we are not immune to the reflection problem in this setting. This is consistent with authors' strategic use of acknowledgements as a signaling mechanism.

Overall, our stylized model and empirical analysis highlight the importance of intellectual collaboration and network effects on the impact academic work can have. The importance of intellectual collaboration adds new insights into the division of labor in academic teams. There is a wide range of activities that are necessary for scientific innovation (Haeussler and Sauermann, 2016). But not all of these need to be performed by co-authors only, i.e. authors in economics can extend the team to outsource activities that do not justify co-authorship. For example, authors test arguments and the scope of their article's contribution while presenting, or they rely on trusted assessors for relevant literature. It is precisely these larger groups that we target at.

A Appendix

A.1 Proof of Proposition 1

Each scientist chooses a level of effort that maximizes (2). The respective first order condition is

$$\beta e_i^* = t_i + \alpha \sum_{j \in N_i} g_{ij} \frac{e_j^* t_j}{p_j} \quad \text{for each } i \in N. \tag{11}$$

If D_p is the diagonal matrix consisting of p_i's as diagonal elements, then (11) can be written in matrix form as

$$\beta \mathbf{e}^* - \alpha \mathbf{e}^* D_{t_p} G = \mathbf{e}^* \left(\beta I - \alpha A\right) = \mathbf{t}$$

Debreu and Herstein (1953, Theorems III^* and III) show that the matrix $\left(\beta I - A\right)$ is well-defined and non-negative, that is $\left(\beta I - A\right) > 0$, whenever $\beta > \mu_1(A)$. They also show that under such conditions, for any pair of vectors $\mathbf{x}$ and $\mathbf{y} \geq 0$, such that $\mathbf{x} = \left(\beta I - A\right)^{-1} \mathbf{y}$, then $\mathbf{x} > 0$. This then implies that

$$\mathbf{e}^* = \frac{1}{\beta} \left(I - \frac{\alpha}{\beta} A\right)^{-1} \mathbf{t} = \frac{1}{\beta} \mathbf{b}(A, \alpha_\beta, \mathbf{t}) \tag{12}$$

is a unique interior equilibrium vector whenever $\beta > \alpha \mu_1(A)$. It remains to show that $\mu_1(A) \geq \frac{t_n}{p_1} \mu_1(G)$. This relation follows from the following lemma.

Lemma 1. *Wang and Xi (1997, Lemma 2) Let $G, H \in \mathbb{R}^{n \times n}$ be positive semidefinite Hermitian, and let $1 \leq i_1 < \cdots < i_k \leq n$. Then*

$$\prod_{\tau=1}^{k} \mu_{i_\tau}(GH) \geq \prod_{\tau=1}^{k} \mu_{i_\tau}(G) \mu_{n-\tau+1}(H) \tag{13}$$

So if $\tau = 1$, then (13) implies that $\mu_1(D_{t_p} G) \geq \mu_1(G) \mu_n(D_{t_p})$. Since $\mu_n(D_{t_p})$ is the least eigenvalue of D_{t_p}, it follows that $\mu_n(D_{t_p}) = \frac{t_n}{p_1}$. Hence $\mu_1(A) \geq \frac{t_n}{p_1} \mu_1(G)$

A.2 Proof of Corollary 1

Considering the case in which all scientists are involved in the same number of projects p, equilibrium relation can be expressed as

$$\beta \mathbf{e}^* - \alpha \frac{t}{p} \mathbf{e}^* G = t \mathbf{1} \tag{14}$$

Now, consider the two networks G and $G' = G + D$, with p and $p' = p + x$ as the respective number of projects. Let $\mathbf{e}^*$ and $\mathbf{e}^{*'}$ be equilibrium configuration corresponding to G and G'

respectively. Assuming that $t_i = t$ for all $i \in N$, then the respective equilibrium conditions are

$$\alpha t\mathbf{e}^* G = p\beta\mathbf{e}^* - pt\mathbf{1} \tag{15}$$

$$\alpha t\mathbf{e}^{*'} G' = p'\beta\mathbf{e}^{*'} - p't\mathbf{1} \tag{16}$$

Let S_e and $S_{e'}$ be the sum of the elements of the vector $\mathbf{e}^*$ and $\mathbf{e}^{*'}$ respectively. Let also $S_{ee'} = \sum_{i\in N} e_i^* e_i^{*'}$ Multiplying both sides of (16) by $\mathbf{e}^{*^T}$, the transpose of $\mathbf{e}^*$, yields

$$\alpha t\mathbf{e}^{*'}(G+D)\mathbf{e}^{*^T} = p'\beta\mathbf{e}^{*'}\mathbf{e}^{*^T} - p't\mathbf{1}\mathbf{e}^{*^T} = p'\beta S_{ee'} - p'tS_e \tag{17}$$

Note that since G is symmetric, $\alpha t\mathbf{e}^* G \equiv \alpha tG\mathbf{e}^{*^T} = p\beta\mathbf{e}^{*^T} - pt\mathbf{1}^T$. Substituting (15) into (17) gives $p\beta S_{ee'} - ptS'_e + \alpha t\mathbf{e}^{*'}D\mathbf{e}^{*^T} = p'\beta S_{ee'} - p'tS_e$. Substituting for $p' = p + x$ then yields

$$S'_e = S_e + \frac{x}{p}S_e - \frac{x}{pt}\beta S_{ee'} + \frac{\alpha}{p}\mathbf{e}^{*'}D\mathbf{e}^{*^T} \tag{18}$$

Hence, $S'_e > S_e$ if and only if $x(\beta S_{ee'} - tS_e) < \alpha t\mathbf{e}^{*'}D\mathbf{e}^{*^T}$.

A.3 Tables

Table 1: Overview of variables used in the different samples of this paper.

Name	N	definition	purpose	variables
full sample	5808	all research articles published in The Journal of Finance, The Review of Financial Studies, the Journal of Financial Economics, the Journal of Financial Intermediation, the Journal of Money, Credit & Banking, and the Journal of Banking and Finance, published in the 1997-2011 period	describe facts on intensive and extensive margin of collaboration in financial economics	number of authors, number of commenters, number of seminars, number of conferences
age sample	2,404	all items from *full sample* that acknowledge at least one conference along with the year in which it was held	regression of publication age on various measures of informal collaboration	publication age, number of commenters, number of seminars, number of conferences
NBER sample	161 (389)	all manuscripts discussed in finance-related NBER summer meetings between 2001 and 2011 that eventually were published. The list of relevant NBER summer institutes is the following: Monetary Economics (15 meetings), Asset Pricing (11), Corporate Finance (11), International Finance & Macroeconomics (11), Risk of Financial Institutions (6), Household Finance (2), Finance & Macro Meeting (1), Asset Marketing/Real Estate (1)	regression of discussant eigenvecotr centrality on citations	number of authors, number of pages, paper age, sum of citation stock of all authors, sum of experience of all authors, sum of experience of authors squared, sum of citations stock of all discussants, sum of experience of discussant all discussants, sum of experience of discussant all discussants squared, eigenvector centrality rank of most eigenvector central discussant, year of publication dummy, journal dummy

Table 1: *(continued)*

Name	N	definition	purpose	variables
commenter sample	1,356	all items from *full sample* that are a) published in the 2009-2011 period and b) list at least one commenter (excluding discussants, editors and referees, including PhD committee members)	regression of commenter centrality on citations and publication status	number of authors, number of pages, number of commenters, number of seminars, number of conferences, top conference dummy, citations of most-cited author, experience of most-experienced author, experience of most-experienced author squared, degree of commenter with highest degree, betweenness centrality of most betweenness central commenter, eigenvector of most eigenvector central commenter, JEL codes dummies, journal dummy

Table 2: Overview of dependent variables used in this paper.

Variable	Definition
Age (*age sample*)	Number of years between year of publication and the year of the oldest conference as listed in the acknowledgement section
Total citations	Number of citations of an article until January 2017, obtained from Scopus (*full sample* and *NBER sample*)
Top publication prob.	Empirical probability that article was published in a top Finance journal (*full sample* and *NBER sample*)

Table 3: Overview of independent variables used in this paper.

Variable	Purpose	Definition
Auth. total experience/ Com. total experience	author controls/ commenter controls	Number of years between publication year minus 1 and the year in which the author's (commenter's) first article was published; sum there of if multi-authored article (for commenters: sum over all acknowledged commenters)
Auth. total experience^2/ Com. total experience^2	author controls/ commenter controls	The square of Authors' (Commenters') total experience; sum thereof if multi-authored article (for commenters: sum over all acknowledged commenters)
Auth. total Euclid/ Com. total Euclid	author controls/ commenter controls	Euclidean Index of an author's (commenter's) publications (square root of sum of squared total citations to each article in a given year) in the year before publication of the article under consideration; sum thereof if multi-authored article (for commenters: sum over all acknowledged commenters)
Auth. total projects/ Com. total projects	author controls/ Commenter controls	Number of co-author adjusted publications of an author (a commenter) in the year before publication and the following year; sum thereof if multi-authored article (for commenters: sum over all acknowledged commenters)
Dis. total experience	discussant controls	Number of years between the year of the discussion and the year in which the discussant's first article was published; sum thereof if article was discussed multiple times
Dis. total experience^2	discussant controls	The square of Dis. total experience; sum thereof if article was discussed multiple times
Dis. total Euclid	discussant controls	Euclidean Index of a discussant's publications (square root of sum of squared total citations to each article in a given year) in the year of the discussion; sum thereof if article was discussed multiple times
Dis. total projects	discussant controls	Number of co-author adjusted publications of a discussant in the year of the discussion and the following year; sum thereof if article was discussed multiple times
Dis. total Bonacich rank	main independent variable	Rank according to Bonacich centrality of the article's discussant, measured in the social network of informal collaboration for the three years before the discussion, where links are weighted by frequency and productivity of the involved researchers; sum thereof if article was discussed multiple times
number of seminars	stylized facts; independent variable	Number of seminars acknowledged in the paper's acknowledgement section
top conference dummy	independent variable	binary variable equaling 1 if the paper has been presented at at least one EFA or AFA annual meeting and/or during an NBER summer institute
age	stylized facts; bibliographic controls	Number of years between year of publication and the year of the first presentation at an NBER summer institute
number of pages	stylized facts; bibliographic controls	Number of pages, according to Web of Science
number of authors	stylized facts; bibliographic controls	Number of authors on a paper

Table 3: Overview of independent variables used in this paper.

Variable	Purpose	Definition
number of commenters	stylized facts	Number of researchers acknowledged in the paper's acknowledgement section (includes PhD advisers, excludes discussants, RAs, editors, referees and industry personnel)
number of conferences	stylized facts	Number of conferences acknowledged in the paper's acknowledgement section
number of discussants	bibliographic controls	Number of discussants of a manuscript at finance-related NBER summer institutes

Table 4: Summary statistics for measures of informal collaboration and age.

Statistic	N	Mean	Median	St. Dev.	Min	Max
# of commenters	2,164	9.80	8	6.68	1	52
# of seminars	1,663	5.19	4	4.24	1	30
# of conferences	2,314	2.84	2	1.93	1	13
Age	2,314	2.94	3	1.49	0	12
Age^2	2,314	10.86	9	11.84	0	144

Notes: # of commenters, # of seminars and *# of conference* is the number of commenters, seminars and conferences acknowledged on a research article. *Age* is the difference between the publication year and the earliest year denoted in conference listings. Only research articles with at least one conference whose name indicates a year considered.

Table 5: Spearman and Pearson correlations for measures of informal collaboration and age.

# of commenters		0.31	0.28	0.09
# of seminars	0.30		0.20	0.13
# of conferences	0.28	0.23		0.18
Age	0.08	0.16	0.14	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *# of commenters, # of seminars* and *# of conference* is the number of commenters, seminars and conferences acknowledged on a research article. *Age* is the difference between the publication year and the earliest year denoted in conference listings. Only research articles with at least one conference whose name indicates a year considered.

Table 6: Results of Negative Binomial regression for measures of informal collaboration on age.

	# of commenters	# of seminars	# of conferences
	(1)	(2)	(3)
Age	0.080^{***}	0.097^{***}	0.221^{***}
	(0.026)	(0.035)	(0.028)
Age^2	-0.005	-0.001	-0.021^{***}
	(0.003)	(0.004)	(0.004)
Constant	2.305^{***}	1.621^{***}	0.760^{***}
	(0.065)	(0.087)	(0.068)
Publication year-fixed effects	*Yes*	*Yes*	*Yes*
Journal-fixed effects	*Yes*	*Yes*	*Yes*
N	2,164	1,663	2,314
Log Likelihood	-6,617.191	-4,168.214	-4,309.229

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects and show the per cent increase in the dependent variable in response to a 1 unit increase in the independent variable, holding all variables at their mean and setting binary variables to 0. *# of commenters*, *# of seminars* and *# of conference* is the number of commenters, seminars and conferences acknowledged on a research article. *Age* is the difference between the publication year and the earliest year denoted in conference listings. Only research articles with considered, where at least one acknowledged conference name indicates the year in which it was hold.

Table 7: Comparison of the network of intellectual collaboration at different years.

Panel A: 1999-2002.

Nodes	(Giant)	Components	Edges	Density	Diameter
3556	3376	37	11758	0.002	14

Panel B: 2003-2005.

Nodes	(Giant)	Components	Edges	Density	Diameter
4518	4385	31	16525	0.0017	13

Panel C: 2006-2008.

Nodes	(Giant)	Components	Edges	Density	Diameter
5717	5527	43	23136	0.0015	14

Panel D: 2009-2011.

Nodes	(Giant)	Components	Edges	Density	Diameter
6997	6750	54	32077	0.0014	15

Notes: Tables present network statistics for selected networks of informal intellectual collaboration, where nodes represent financial economists that have collaborated informally on research articles. Each network was inferred from articles published in year t, $t-1$ and $t-2$. *Nodes* is the number of researchers in the network; *(Giant)* is the number of researchers in the giant component, the largest connected component; *Components* is the number of distinct network components; *Edges* is the number of edges/ties connecting the nodes; *Density* is the share of realized to potential paths; *Diameter* is the longest of all shortest paths between all nodes.

Table 8: Published research articles presented in financial NBER summer institutes.

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	total	share
AMRE	4 (5)	5 (6)										9 (11)	82%
AP	5 (8)	8 (9)	7 (8)	5 (8)	8 (9)	8 (9)	7 (9)	7 (7)	7 (8)	8 (9)	10 (12)	80 (96)	83%
CF	4 (6)	6 (9)	8 (10)	10 (13)	12 (14)	13 (14)	10 (13)	16 (18)	5 (6)	13 (14)	5 (6)	102 (123)	83%
EFEL	15 (19)	12 (15)	11 (15)	4 (11)	9 (11)	12 (15)	8 (11)	8 (10)	8 (12)	5 (10)	6 (12)	98 (141)	70%
HF										4 (6)	3 (7)	7 (13)	54%
IFM	7 (8)	10 (12)	8 (11)	9 (12)	5 (8)	7 (10)	7 (10)	8 (12)	11 (14)	9 (15)	8 (12)	89 (124)	72%
ME											4 (6)	4 (6)	67%
RISK						3 (3)	1 (2)	2 (3)	12 (15)	6 (7)	4 (6)	28 (36)	78%
total	35 (46)	41 (51)	34 (44)	28 (44)	34 (42)	43 (51)	33 (45)	41 (50)	43 (55)	45 (61)	40 (61)	417 (550)	
share	76%	80%	77%	64%	81%	84%	73%	82%	78%	74%	66%		76%

Notes: Table lists the number of research articles in the NBER sample by NBER summer institute (year + NBER group). The NBER groups are Asset Marketing/Real Estate (AMRE), Asset Pricing (AP), Corporate Finance (CF), Capital Markets in the Economy (EFEL), Household Finance (HF), International Finance & Macroeconomics (IFM), Monetary Economics (ME), Risk of Financial Institutions (RISK). Numbers in parenthesis indicate the number of discussed presentations during this summer institutes; the difference are thus articles that were not (yet) published in Scopus-indexed journals. For example, during the 2001 AP summer institute, 8 manuscripts were discussed of which 7 were eventually published in a Scopus-indexed journal. Column "share" indicates the share of manuscripts discussed per NBER group resp. year which were eventually published. Note: Includes double counts.

Table 9: Most Bonacich central discussants in the network of informal intellectual collaboration in different years.

	2002	2005	2008	2011
1	Wurgler, Jeffrey (2)	Wurgler, Jeffrey (2)	Campbell, John Y. (5)	Sufi, Amir (1)
2	Stein, Jeremy C. (3)	Greenwood, Robin M. (1)	Stein, Jeremy C. (3)	Edmans, Alex (1)
3	Zingales, Luigi (4)	Stein, Jeremy C. (3)	Baker, Malcolm P. (7)	Greenwood, Robin M. (1)
4	Campbell, John Y. (5)	Baker, Malcolm P. (7)	Levine, Ross L. (2)	Stein, Jeremy C. (3)
5	Baker, Malcolm P. (7)	Jenter, Dirk (2)	Wolfenzon, Daniel (1)	Acharya, Viral V. (4)
6	Johnson, Simon (4)	Zingales, Luigi (4)	French, Kenneth R. (2)	Roberts, Michael R. (2)
7	Wolfenzon, Daniel (1)	Moskowitz, Tobias J. (1)	Claessens, Stijn (1)	Carlin, Bruce I. (1)
8	Sapienza, Paola (2)	Campbell, John Y. (5)	Gomes, Francisco J. (1)	Zingales, Luigi (4)
9	Levine, Ross L. (2)	Wolfenzon, Daniel (1)	Wurgler, Jeffrey (2)	Diamond, Douglas W. (1)
10	Lamont, Owen A. (4)	Johnson, Simon (4)	Zingales, Luigi (4)	Rauh, Joshua D. (1)
11	Barberis, Nicholas C. (1)	Sapienza, Paola (2)	Wachter, Jessica A. (1)	Baker, Malcolm P. (7)
12	Glaeser, Edward (1)	Diamond, Douglas W. (1)	Titman, Sheridan D. (1)	Purnanandam, Amiyatosh K. (1)
13	Allen, Franklin (3)	Barberis, Nicholas C. (1)	Daniel, Kent D. (3)	Metrick, Andrew (2)
14	Xiong, Wei (1)	Lerner, Josh (2)	Strahan, Philip E. (5)	Petersen, Mitchell A. (4)
15	Titman, Sheridan D. (1)	Levine, Ross L. (2)	Yogo, Motohiro (2)	Cohen, Lauren H. (2)
16	Scharfstein, David S. (4)	Mian, Atif (5)	Shiller, Robert J. (1)	Kashyap, Anil K. (1)
17	Cornelli, Francesca (1)	Schoar, Antoinette (4)	Lettau, Martin (1)	Moskowitz, Tobias J. (1)
18	Singleton, Kenneth J. (1)	MÃÿrck, Randall K. (5)	Robinson, David T. (1)	Goldstein, Itay (1)
19	Metrick, Andrew (2)	Mullainathan, Sendhil (1)	MÃÿrck, Randall K. (5)	Sapienza, Paola (2)
20	Daniel, Kent D. (3)	Glaeser, Edward (1)	Jenter, Dirk (2)	Wurgler, Jeffrey (2)
21	French, Kenneth R. (2)	Pedersen, Lasse Heje (1)	Coval, Joshua D. (2)	Landier, Augustin (2)
22	Claessens, Stijn (1)	Froot, Kenneth A. (1)	Moskowitz, Tobias J. (1)	Mian, Atif (5)
23	Stambaugh, Robert F. (2)	Malmendier, Ulrike (1)	Johnson, Simon (4)	Scharfstein, David S. (4)
24	Desai, Mihir A. (1)	Bekaert, Geert (1)	Bekaert, Geert (1)	Campbell, John Y. (5)
25	Gromb, Denis (1)	Scharfstein, David S. (4)	Xiong, Wei (1)	Servaes, Henri (1)

Notes: Table lists most Bonacich central discussants in the network of informal intellectual collaboration for the years 2002, 2005, 2008 and 2011. Bonacich centrality is computed in the largest component of the social network of informal intellectual collaboration in the respective year according to equation (7). Numbers in parentheses indicate the number of discussions in our sample, i.e. of manuscripts discussed in finance-related NBER summer institutes since 2001 that were published in Scopus-index journals before March 2016.

Table 10: Summary statistics for all continuous variables used in the NBER sample.

	N	Mean	Median	Std.Dev.	Min	Max
Impact measure						
Total citations	161	85.0	49	135.64	1	1023
Article characteristics						
# of authors	161	2.4	2	0.83	1	4
# of pages	161	34.8	35	9.78	8	64
# of discussants	161	1.0	1	0.22	1	2
Age	161	3.3	3	1.48	0	9
Age^2	161	12.8	9	11.97	0	81
Author characteristics						
Auth. total Euclid	147	490.4	204	1252.71	0	12319
Auth. total projects	148	21.3	8	40.49	0	288
Auth. total experience	161	23.8	19	18.87	−1	89
Auth. total experience^2	161	919.5	361	1430.55	0	7921
Discussant characteristics						
Dis. total Euclid	158	187.6	93	226.47	0	1002
Dis. total projects	143	2.1	2	1.58	0	8
Dis. total experience	161	12.1	10	8.57	−5	35
Dis. total experience^2	161	219.7	100	262.85	0	1225
Network measure						
Dis. total Bonacich rank	144	459.4	174	626.05	3	3626

Notes: Total citations is the count of citations since publication. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience^2* is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants.

Table 11: Spearman and Pearson correlations for all continuous variables in the NBER sample.

Impact measure													
Total citations		−0.12	0.00	0.04	−0.44	−0.44	0.05	0.11	−0.09	0.04	0.08	0.07	−0.08
Article characteristics													
# of authors	−0.09		−0.02	−0.03	0.06	0.06	0.62	0.58	0.72	0.09	−0.02	0.01	−0.11
# of pages	0.08	−0.04		−0.08	0.11	0.11	0.10	0.04	−0.01	−0.04	0.00	−0.09	0.07
# of discussants	−0.05	−0.03	−0.07		0.00	0.00	−0.08	0.07	−0.04	0.01	0.07	0.11	0.24
Age	−0.26	0.06	0.01	−0.02		1.00	0.00	−0.03	0.10	−0.02	−0.01	−0.03	0.06
Age^2	−0.24	0.04	−0.03	−0.03	0.96		0.00	−0.03	0.10	−0.02	−0.01	−0.03	0.06
Author characteristics													
Auth. total Euclid	−0.03	0.22	−0.12	0.22	−0.11	−0.10		0.54	0.88	0.12	0.01	0.10	0.00
Auth. total projects	−0.01	0.27	−0.01	0.14	−0.08	−0.10	0.23		0.55	0.10	−0.01	0.12	−0.07
Auth. total experience	−0.10	0.68	−0.04	−0.03	0.13	0.13	0.44	0.28		0.11	0.06	0.06	−0.01
Discussant characteristics													
Dis. total Euclid	−0.01	0.12	−0.01	0.01	0.03	0.04	0.09	0.15	0.19		0.26	0.83	−0.06
Dis. total projects	0.15	0.02	0.07	0.05	−0.03	−0.03	0.15	0.09	0.19	0.34		0.20	−0.09
Dis. total experience	0.02	0.01	−0.11	0.21	0.02	0.03	0.24	0.16	0.15	0.65	0.25		0.20
Network measure													
Dis. total Bonacich rank	−0.18	−0.04	−0.04	0.25	0.05	0.06	0.08	0.08	0.08	−0.10	−0.14	0.26	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total citations* is the count of citations since citation. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants. Only articles published before January 2017 considered.

Table 12: Results of Negative Binomial regression for discussant characteristics of published manuscripts presented at finance-related NBER sumemr institutes.

	Dis. Euclid	Dis. projects	Dis. experience	Dis. Bonacich rank	Dis. betweenness rank
Auth. total Euclid	−0.00003	0.00004	0.00003	0.0001*	0.0003***
	$p = 0.743$	$p = 0.412$	$p = 0.478$	$p = 0.080$	$p = 0.002$
Auth. total projects	0.003	0.002	0.006	0.006	−0.003
	$p = 0.905$	$p = 0.892$	$p = 0.652$	$p = 0.803$	$p = 0.912$
Auth. total experience	−0.011	0.009	−0.002	−0.016	−0.004
	$p = 0.579$	$p = 0.574$	$p = 0.871$	$p = 0.480$	$p = 0.843$
Auth. total experience^2	0.001	−0.0001	0.0002	0.001	0.0002
	$p = 0.277$	$p = 0.688$	$p = 0.568$	$p = 0.236$	$p = 0.664$
# of authors	−0.086	−0.023	−0.093*	−0.146	−0.134
	$p = 0.321$	$p = 0.729$	$p = 0.075$	$p = 0.117$	$p = 0.137$
Constant	5.540***	0.528*	2.563***	6.233***	6.943***
	$p = 0.000$	$p = 0.100$	$p = 0.000$	$p = 0.000$	$p = 0.000$
Discussion year-fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
NBER group-fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	360	346	357	317	317
Log Likelihood	-2,203.212	-586.318	-1,194.625	-2,359.744	-2,517.309

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects and show the per cent increase in the dependent variable in response to a 1 unit increase in the independent variable, holding all variables at their mean and setting binary variables to 0. *Dis. Euclid* is the discussants Euclidean index as defined in (8) in the year of the discussion. *Dis. projects* is the number of current projects of the discussant in the year of the discussion. *Dis. experience* is the number of years between first publication and the year of the discussion. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year of the the discussion, summed over all authors. *Auth. total projects* is the number of other publications in the year of the discussion and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year of the discussion and the publication year of the first article, summed over all authors. *Auth. total experience*^{2} is its square. *# of authors* is the size of the author group. Only discussants of presentations at finance-related NBER summer institutes, eventually resulted in publication, considered.

Table 13: Results of Negative Binomial regression for citation count, *NBER sample*.

	(1)	(2)	(3)	(4)	(5)
Auth. total Euclid	0.00000	0.00001	0.00001	0.0001	0.00002
	$p = 0.978$	$p = 0.914$	$p = 0.935$	$p = 0.311$	$p = 0.807$
Auth. total projects	0.001	0.001	0.001	0.004^{***}	0.005^{***}
	$p = 0.343$	$p = 0.719$	$p = 0.553$	$p = 0.006$	$p = 0.001$
Auth. total experience	0.035^{***}	0.034^{***}	0.028^{***}	0.013	0.008
	$p = 0.002$	$p = 0.002$	$p = 0.009$	$p = 0.153$	$p = 0.325$
Auth. total experience^2	-0.0004^{***}	-0.0004^{***}	-0.0003^{**}	-0.0002^{*}	-0.0001
	$p = 0.002$	$p = 0.001$	$p = 0.011$	$p = 0.052$	$p = 0.600$
Dis. total Euclid		0.0001	-0.00001	0.0002	0.0002
		$p = 0.687$	$p = 0.979$	$p = 0.505$	$p = 0.385$
Dis. total projects		0.038	0.016	-0.017	0.019
		$p = 0.315$	$p = 0.675$	$p = 0.573$	$p = 0.519$
Dis. total experience		-0.003	-0.011	0.028^{**}	0.030^{**}
		$p = 0.894$	$p = 0.641$	$p = 0.046$	$p = 0.024$
Dis. total experience^2		0.001	0.001	-0.001	-0.001^{**}
		$p = 0.509$	$p = 0.225$	$p = 0.116$	$p = 0.021$
Dis. total Bonacich rank			-0.0003^{**}	-0.0002^{**}	-0.0001^{*}
			$p = 0.016$	$p = 0.048$	$p = 0.090$
Constant	4.550^{***}	4.598^{***}	5.019^{***}	4.467^{***}	4.319^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$
Sample	$NBER$	$NBER$	$NBER$	$NBER(+econ)$	$NBER(+all)$
Article characteristics	Yes	Yes	Yes	Yes	Yes
NBER group dummies	Yes	Yes	Yes	Yes	Yes
Journal-FE	Yes	Yes	Yes	Yes	Yes
Publication year-FE	Yes	Yes	Yes	Yes	Yes
N	161	161	161	249	389
Log Likelihood	-767.502	-764.554	-761.695	-1,208.431	-1,848.234

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects and show the per cent increase in the citation count in response to a 1 unit increase in the independent variable, holding all variables at their mean and setting binary variables to 0. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants. *Article characteristics* include *# of authors*, *# of discussants # of pages*, *age*, and *age*^{2} being the count of authors, the number of discussants form finance-related NBER summer institutes, the count of pages, and the number of years passed between the first presentation at an NBER summer institute and the publication year, and its square. Journal-fixed effects are relative to The Journal of Finance.

Table 14: Summary statistics for all continuous variables used in the extended NBER sample.

	N	Mean	Median	Std.Dev.	Min	Max
Impact measure						
Top publication prob.	389	0.4	0	0.49	0	1
Article characteristics						
# of authors	389	2.3	2	0.79	1	5
# of pages	389	32.1	33	11.43	1	73
# of discussants	389	1.1	1	0.28	1	2
Age	389	3.5	3	1.89	0	11
Age^2	389	15.8	9	17.15	0	121
Author characteristics						
Auth. total Euclid	353	404.0	189	884.23	0	12319
Auth. total projects	354	22.1	8	37.69	0	288
Auth. total experience	389	24.8	22	17.28	−1	91
Auth. total experience^2	389	910.5	484	1234.33	0	8281
Discussant characteristics						
Dis. total Euclid	380	177.7	88	250.90	0	1592
Dis. total projects	341	2.1	2	1.68	0	11
Dis. total experience	389	12.8	10	9.68	−5	54
Dis. total experience^2	389	256.3	100	350.50	0	2916
Network measure						
Dis. total Bonacich rank	313	837.2	377	1015.47	2	5223

Notes: Top publication prob. is the empirical probability that the article was published in one of the top three Finance journals The Journal of Finance, The Review of Financial Studies and the Journal of Financial Economics. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants.

Table 15: Spearman and Pearson correlations for all continuous variables in the NBER sample.

Impact measure													
Top publication prob.		0.11	0.22	−0.10	−0.08	−0.08	0.01	−0.04	−0.09	0.06	0.07	−0.03	−0.35
Article characteristics													
# of authors	0.10		0.06	−0.02	0.03	0.03	0.52	0.54	0.61	−0.01	−0.01	−0.05	−0.14
# of pages	0.20	0.04		−0.08	−0.05	−0.05	0.14	0.16	−0.03	0.06	−0.01	−0.03	−0.16
# of discussants	−0.10	−0.02	−0.09		0.06	0.06	−0.04	0.08	0.00	0.15	0.21	0.26	0.28
Age	−0.11	0.02	−0.09	0.05		1.00	0.06	−0.07	0.17	−0.07	−0.06	−0.01	0.08
Age^2	−0.15	0.02	−0.09	0.04	0.95		0.06	−0.07	0.17	−0.07	−0.06	−0.01	0.08
Author characteristics													
Auth. total Euclid	0.08	0.21	0.00	0.09	−0.01	0.01		0.52	0.82	0.13	0.00	0.11	−0.01
Auth. total projects	−0.02	0.30	0.13	0.11	−0.06	−0.04	0.26		0.47	0.08	0.04	0.10	−0.04
Auth. total experience	−0.05	0.61	−0.03	0.01	0.17	0.17	0.44	0.31		0.07	0.01	0.08	0.02
Discussant characteristics													
Dis. total Euclid	0.03	0.01	0.04	0.17	−0.01	−0.02	0.11	0.03	0.12		0.24	0.82	−0.09
Dis. total projects	0.04	−0.01	0.03	0.18	−0.08	−0.06	0.10	0.10	0.07	0.32		0.21	−0.03
Dis. total experience	−0.06	−0.05	−0.04	0.35	0.02	0.02	0.19	0.08	0.10	0.65	0.30		0.19
Network measure													
Dis. total Bonacich rank	−0.34	−0.09	−0.18	0.28	0.06	0.08	−0.02	0.01	0.01	−0.13	−0.01	0.18	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Top publication prob.* is the empirical probability that the article was published in one of the top three Finance journals The Journal of Finance, The Review of Financial Studies and the Journal of Financial Economics. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants. Only articles published before January 2017 considered.

Table 16: Results of Logit regression for probability of publishing in top journals, *NBER sample.*

	(1)	(2)	(3)
Auth. total Euclid	0.001^{**}	0.001^{**}	0.0005^{*}
	$p = 0.030$	$p = 0.033$	$p = 0.051$
Auth. total projects	-0.003	-0.0003	-0.0003
	$p = 0.368$	$p = 0.937$	$p = 0.948$
Auth. total experience	-0.088^{***}	-0.095^{***}	-0.099^{***}
	$p = 0.0003$	$p = 0.0004$	$p = 0.0003$
Auth. total experience^2	0.001^{***}	0.001^{***}	0.001^{***}
	$p = 0.004$	$p = 0.009$	$p = 0.004$
Dis. total Euclid		0.001	-0.0001
		$p = 0.220$	$p = 0.945$
Dis. total projects		0.112	0.107
		$p = 0.205$	$p = 0.241$
Dis. total experience		0.032	0.049
		$p = 0.527$	$p = 0.331$
Dis. total experience^2		-0.002	-0.002
		$p = 0.216$	$p = 0.248$
Dis. total Bonacich rank			-0.001^{***}
			$p = 0.001$
Constant	-2.347^{**}	-2.502^{*}	-2.318^{*}
	$p = 0.030$	$p = 0.053$	$p = 0.083$
Other collaboration	*Yes*	*Yes*	*Yes*
Publication year-fixed effects	*Yes*	*Yes*	*Yes*
N	353	310	310
Log Likelihood	-198.560	-169.272	-162.834

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects and show the per cent increase in the empirical probability that article was published in a top journal in response to a 1 unit increase in the independent variable, holding all variables at their mean and setting binary variables to 0. Top journal refers to The Journal of Finance, The Review of Financial Studies, and the Journal of Financial Economics. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all discussants. *Article characteristics* include *# of authors, # of discussants # of pages, age*, and *age*^{2} being the count of authors, the number of discussants form finance-related NBER summer institutes, the count of pages, and the number of years passed between the first presentation at an NBER summer institute and the publication year, and its square.

Table 17: Summary statistics for all continuous variables used in the *commenter sample*.

	N	Mean	Median	Std.Dev.	Min	Max
Impact measure						
Total citations	4406	73.8	37	118.46	0	2163
Top publication prob.	4406	0.6	1	0.49	0	1
Article characteristics						
# of authors	4406	2.2	2	0.84	1	6
# of pages	4406	26.7	26	10.29	3	80
Informal Collaboration						
# of commenters	4406	9.0	8	6.47	1	58
# of seminars	2800	5.2	4	4.21	1	30
# of conferences	3067	2.8	2	1.97	1	23
Top conference	4406	0.1	0	0.26	0	1
Author characteristics						
Auth. total Euclid	4406	214.5	78	432.48	0	8991
Auth. total projects	4406	3.7	3	3.06	0	50
Auth. total experience	4406	20.7	18	16.55	0	104
Auth. total experience^2	4406	702.9	324	1059.04	0	10816
Commenter characteristics						
Com. total Euclid	4383	1566.7	849	2086.30	0	21231
Com. total projects	4323	12.0	9	10.69	0	118
Com. total experience	4393	116.3	93	93.08	−5	774
Com. total experience^2	4393	22176.6	8649	39733.18	0	599076
Network measure						
Com. total Bonacich rank	4406	14340.5	10494	13363.29	2	127697

Notes: Total citations is the count of citations since publication. *Top publication prob.* is the probability that the article was published in one of the three top journals (JF, RFS, JFE). *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience^2* is its square. *# of commenters, # of seminars* and *# of conferences* are the count of acknowledged commenters resp. seminars resp. conferences. *Top conference* is a dummy variable equaling 1 when the article has been presented at least at one AFA or EFA conference or at an NBER meeting.*Com. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all acknowledged commenters. Only articles with at least one commenter considered.

Table 18: Spearman and Pearson correlations for all continuous variables in the *commenter sample*.

Impact measure															
Total citations		1.00	0.46	0.37	0.11	0.05	0.46	0.18	0.33	0.17	0.37	0.02	0.31	0.30	0.18
Top publication prob.	0.29		0.12	0.23	0.05	0.21	1.00	0.18	−0.17	0.06	0.56	0.33	0.14	0.32	0.17
Article characteristics															
# of authors	0.07	0.09		0.20	0.22	0.21	0.09	0.24	0.12	0.56	0.04	0.27	0.03	1.00	0.18
# of pages	0.24	0.54	0.04		0.12	0.17	0.56	−0.01	0.09	0.45	1.00	0.30	−0.04	0.21	0.31
Informal collaboration															
# of commenters	0.15	0.32	0.00	0.28		−0.21	0.36	0.04	1.00	0.65	0.32	−0.10	−0.12	0.14	0.21
# of seminars	0.15	0.30	0.01	0.26	0.31		0.34	0.42	0.04	0.05	0.30	0.23	0.71	0.17	1.00
# of conferences	0.05	0.17	0.16	0.17	0.31	0.25		0.26	−0.01	−0.01	0.18	0.36	0.78	0.02	0.19
Top conference	0.00	0.18	0.05	0.10	0.15	0.14	0.21		0.01	0.01	0.11	−0.01	0.88	0.02	0.15
Author characteristics															
Auth. total Euclid	0.16	0.18	0.29	0.10	0.05	0.11	0.10	0.09		−0.06	0.17	0.32	0.57	0.35	0.11
Auth. total projects	0.07	−0.02	0.36	0.00	0.00	0.02	0.10	0.04	0.33		0.00	1.00	0.20	0.29	0.09
Auth. total experience	0.03	0.03	0.63	0.01	−0.09	0.04	0.07	0.02	0.50	0.39		0.32	0.34	0.32	0.27
Commenter characteristics															
Com. total Euclid	0.09	0.31	0.05	0.22	0.63	0.31	0.24	0.22	0.21	0.09	0.02		0.01	0.00	0.27
Com. total projects	0.15	0.22	0.00	0.24	0.78	0.27	0.27	0.16	0.12	0.11	−0.06	0.67		0.11	0.29
Com. total experience	0.09	0.30	0.02	0.26	0.89	0.32	0.29	0.18	0.10	0.04	−0.03	0.74	0.79		0.19
Network measure															
Com. total Bonacich rank	−0.15	−0.16	−0.02	−0.11	0.54	0.00	0.18	0.01	−0.14	−0.03	−0.09	0.11	0.34	0.42	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total citations* is the count of citations since publication. *Top publication prob.* is the probability that the article was published in one of the three top journals (JF, RFS, JFE). *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. Same logic applies to *Com.* variables. *Auth. total experience^2* is its square. *# of commenters*, *# of seminars* and *# of conferences* are the count of acknowledged commenters resp. seminars resp. conferences. *Top conference* is a dummy variable equaling 1 when the article has been presented at least at one AFA or EFA conference or at an NBER meeting. *Com. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration according to equation (7) in the year of the discussion, summed over all acknowledged commenters. Only articles with at least one commenter considered.

Table 19: Results of Negative Binomial regression for citation count, commenter sample.

	(1)	(2)	(3)
# of authors	0.078***	0.079***	0.084***
	p = 0.0005	p = 0.0003	p = 0.0002
# of pages	0.010***	0.008***	0.008***
	p = 0.00000	p = 0.00001	p = 0.00002
# of commenters	0.016***	0.026***	0.040***
	p = 0.000	p = 0.00000	p = 0.000
# of seminars	0.015***	0.012***	0.011***
	p = 0.0002	p = 0.002	p = 0.006
# of conferences	0.012	0.014*	0.015**
	p = 0.131	p = 0.058	p = 0.050
Top conference	0.085	0.073	0.051
	p = 0.138	p = 0.194	p = 0.363
Auth. total Euclid	0.001***	0.0005***	0.0004***
	p = 0.000	p = 0.000	p = 0.000
Auth. total projects	0.043***	0.035***	0.034***
	p = 0.000	p = 0.000	p = 0.000
Auth. total experience	−0.002	0.0003	0.0001
	p = 0.380	p = 0.912	p = 0.955
Auth. total experience^2	−0.0001	−0.0001*	−0.0001*
	p = 0.107	p = 0.054	p = 0.067
Com. total Euclid		0.0001***	0.00004***
		p = 0.00000	p = 0.001
Com. total projects		0.016***	0.016***
		p = 0.000	p = 0.000
Com. total experience		−0.002***	−0.002***
		p = 0.00002	p = 0.00004
Com. total experience^2		−0.00000***	−0.00000***
		p = 0.0004	p = 0.0001
Com. total Bonacich rank			−0.00001***
			p = 0.00003
Constant	3.430***	3.369***	3.401***
	p = 0.000	p = 0.000	p = 0.000
JEL dummies	*Yes*	*Yes*	*Yes*
Journal-fixed effects	*Yes*	*Yes*	*Yes*
Publication year-fixed effects	*Yes*	*Yes*	*Yes*
N	4,406	4,406	4,406
Log Likelihood	-22,123.230	-22,058.960	-22,049.580

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects. *# of authors* and *# of pages* is the simple count of authors and pages, respectively. *# of commenters*, *# of seminars* and *# of conferences* are the count of acknowledged commenters resp. seminars resp. conferences. *Top conference* equals 1 if the article has been presented at least at one AFA or EFA conference or at an NBER meeting and is 0 otherwise. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Com.* variables. *Com. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration in the year of the discussion (equation (7)), summed over all acknowledged commenters.

Table 20: Results of Logit regression for probability of publishing in top-3 finance journals, *commenter sample.*

	(1)	(2)	(3)
# of authors	0.411^{***}	0.374^{***}	0.383^{***}
	$p = 0.000$	$p = 0.00000$	$p = 0.00000$
# of pages	0.141^{***}	0.140^{***}	0.133^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.000$
# of commenters	0.067^{***}	0.130^{***}	0.388^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.000$
# of seminars	0.190^{***}	0.177^{***}	0.149^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.000$
# of conferences	-0.080^{***}	-0.068^{***}	-0.080^{***}
	$p = 0.002$	$p = 0.010$	$p = 0.005$
Top conference	1.458^{***}	1.401^{***}	1.084^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.00002$
Auth. total Euclid	0.002^{***}	0.002^{***}	0.001^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.00002$
Auth. total projects	-0.091^{***}	-0.073^{***}	-0.066^{***}
	$p = 0.00000$	$p = 0.00002$	$p = 0.0003$
Auth. total experience	-0.011	-0.007	-0.007
	$p = 0.163$	$p = 0.376$	$p = 0.355$
Auth. total experience^2	-0.0002	-0.0002^{*}	-0.0001
	$p = 0.103$	$p = 0.079$	$p = 0.343$
Com. total Euclid		0.001^{***}	0.0002^{***}
		$p = 0.000$	$p = 0.001$
Com. total projects		-0.061^{***}	-0.069^{***}
		$p = 0.000$	$p = 0.000$
Com. total experience		-0.0004	-0.002
		$p = 0.836$	$p = 0.432$
Com. total experience^2		-0.00002^{***}	-0.00001^{***}
		$p = 0.00001$	$p = 0.0004$
Com. total Bonacich rank			-0.0001^{***}
			$p = 0.000$
Constant	-4.463^{***}	-4.792^{***}	-3.987^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.000$
JEL dummies	*Yes*	*Yes*	*Yes*
Publication year-fixed effects	*Yes*	*Yes*	*Yes*
N	4,406	4,406	4,406
Log Likelihood	-1,848.205	-1,761.274	-1,597.371

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects. Dependent variable equals 1 if the article has been published in one of the three top journals (JF, JFE, RFS) and 0 otherwise. *# of authors* and *# of pages* is the simple count of authors and pages, respectively. *# of commenters*, *# of seminars* and *# of conferences* are the count of acknowledged commenters resp. seminars resp. conferences. *Top conference* is a dummy variable equaling 1 when the article has been presented at least at one AFA or EFA conference or at an NBER meeting. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Similar logic applies to *Com.* variables. *Com. total Bonacich rank* is the rank according to Bonacich centrality measured in the giant component of the network of informal intellectual collaboration in the year of the discussion (equation (7)), summed over all acknowledged commenters.

Table 21: Results of Negative Binomial regressions for count of citations with different centrality definitions, *NBER sample*.

	(1)	(2)	(3)	(4)	(5)
Auth. total Euclid	0.00001	0.00001	0.00001	0.00000	0.00001
	$p = 0.928$	$p = 0.910$	$p = 0.911$	$p = 0.940$	$p = 0.886$
Auth. total projects	0.001	0.001	0.001	0.001	0.001
	$p = 0.599$	$p = 0.591$	$p = 0.723$	$p = 0.582$	$p = 0.552$
Auth. total experience	0.031^{***}	0.031^{***}	0.034^{***}	0.029^{***}	0.030^{***}
	$p = 0.004$	$p = 0.004$	$p = 0.002$	$p = 0.006$	$p = 0.005$
Auth. total experience^2	-0.0003^{***}	-0.0004^{***}	-0.0004^{***}	-0.0003^{***}	-0.0003^{***}
	$p = 0.005$	$p = 0.004$	$p = 0.001$	$p = 0.007$	$p = 0.006$
Dis. total Euclid	0.00004	0.00003	0.0001	0.00004	0.00004
	$p = 0.910$	$p = 0.932$	$p = 0.684$	$p = 0.898$	$p = 0.905$
Dis. total projects	0.025	0.028	0.038	0.017	0.023
	$p = 0.495$	$p = 0.451$	$p = 0.324$	$p = 0.653$	$p = 0.540$
Dis. total experience	−0.007	−0.009	−0.003	0.002	−0.004
	$p = 0.760$	$p = 0.697$	$p = 0.894$	$p = 0.944$	$p = 0.866$
Dis. total experience^2	0.001	0.001	0.001	0.001	0.001
	$p = 0.295$	$p = 0.272$	$p = 0.512$	$p = 0.434$	$p = 0.329$
Dis. total Bonacich rank (95%)	-0.0003^{**}				
	$p = 0.027$				
Dis. total Bonacich rank (90%)		-0.0003^{**}			
		$p = 0.031$			
Dis. total Bonacich rank unweighted			0.00000		
			$p = 0.976$		
Dis. total eigenvector rank				-0.0003^{***}	
				$p = 0.008$	
Dis. total Bonacich rank (comb.)					-0.0003^{**}
					$p = 0.011$
Constant	4.804^{***}	4.780^{***}	4.598^{***}	4.878^{***}	4.874^{***}
	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$
Article characteristics	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
NBER group dummies	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Journal-FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Publication year-FE	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
N	161	161	161	161	161
Log Likelihood	-762.100	-762.154	-764.553	-761.341	-761.462

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects. *Auth. total Euclid* is the author's Euclidean index as defined in equation (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total Bonacich rank (95%)* (*Dis. total Bonacich rank (90%)*) is the rank according to Bonacich centrality with an attenuation factor of 95% (90%) of the network's leading Eigenvalue. *Dis. total Bonacich rank unweighted* is the rank according to Bonacich centrality without link weights, summed over all discussants. *Dis. total eigenvector rank* is the rank according to Eigenvector centrality, summed over all discussants. *Dis. total Bonacich rank (comb.)* is the rank according to Bonacich centrality measured in the network of intellectual collaboration (including co-author ties), summed over all discussants. All network variables measured in the giant component of the respective network in the year of the discussion. *Article characteristics* include *# of authors*, *# of discussants # of pages*, *age*, and *age*^{2} being the count of authors, the number of discussants from finance-related NBER summer institutes, the count of pages, and the number of years passed between the first presentation at an NBER summer institute and the publication year, and its square.

Table 22: Results of Negative Binomial regression for count of citations with betweenness centrality, *NBER sample*.

	(1)
Auth. total Euclid	0.00001
	$p = 0.884$
Auth. total projects	0.001
	$p = 0.689$
Auth. total experience	0.035^{***}
	$p = 0.002$
Auth. total experience^2	-0.0004^{***}
	$p = 0.001$
Dis. total Euclid	0.0001
	$p = 0.779$
Dis. total projects	0.039
	$p = 0.298$
Dis. total experience	-0.008
	$p = 0.760$
Dis. total experience^2	0.001
	$p = 0.447$
Dis. total betweenness	0.063
	$p = 0.299$
Constant	4.558^{***}
	$p = 0.000$
Article characteristics	*Yes*
NBER group dummies	*Yes*
Journal-FE	*Yes*
Publication year-FE	*Yes*
N	161
Log Likelihood	-764.048

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Reported coefficients are marginal effects and show the per cent increase in the empirical probability that article was published in a top journal in response to a 1 unit increase in the independent variable, holding all variables at their mean and setting binary variables to 0. Top journal refers to The Journal of Finance, The Review of Financial Studies, and the Journal of Financial Economics. *Auth. total Euclid* is the author's Euclidean index as defined in (8) in the year before publication, summed over all authors. *Auth. total projects* is the number of other publications in the year of publication and the following year, divided by the number of co-authors, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of the first article of all authors. *Auth. total experience*^{2} is its square. Same logic applies to *Dis.* variables. *Dis. total betweenness* is the normalized betweenness centrality score according to equation (10) measured in the giant component of the network of informal intellectual collaboration in the year of the discussion without link weights, summed over all discussants. *Article characteristics* include *# of authors*, *# of discussants # of pages*, *age*, and *age*^{2} being the count of authors, the number of discussants form finance-related NBER summer institutes, the count of pages, and the number of years passed between the first presentation at an NBER summer institute and the publication year, and its square.

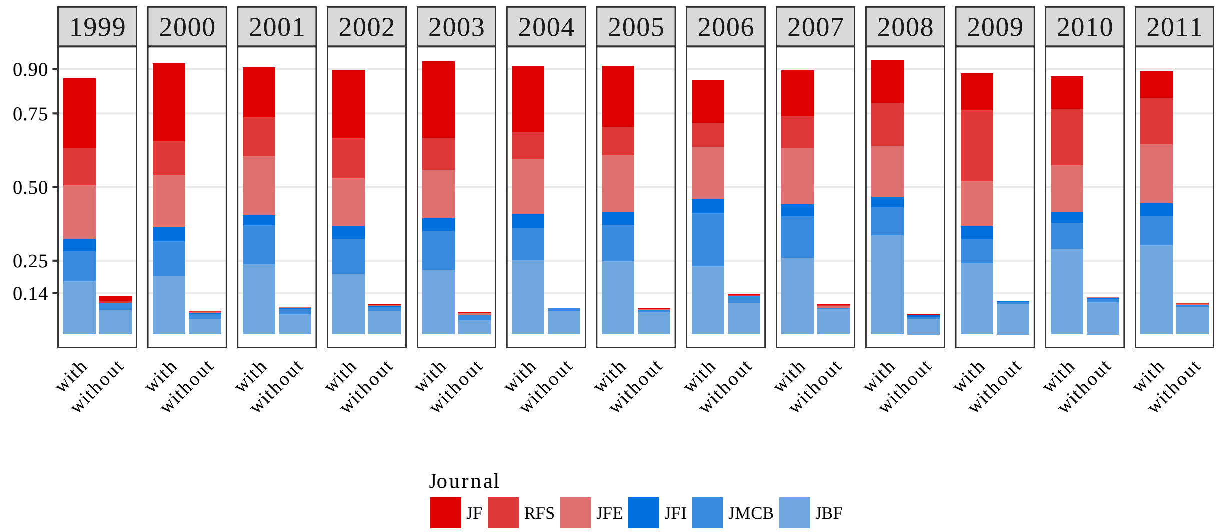
Table 23: Regression results replicating Laband and Tollison (2000) and Brown (2005)

	Five year citations			Total citations
	OLS			*negative binomial*
	(1)	(2)	(3)	(4)
Auth. total projects	1.074***	0.920***	0.959***	
	(0.135)	(0.135)	(0.135)	
# of authors				0.151***
				(0.017)
# of pages	0.455***	0.488***	0.445***	0.012***
	(0.042)	(0.041)	(0.042)	(0.002)
# of commenters	0.680***		0.368***	0.017***
	(0.066)		(0.078)	(0.002)
Com. total citations		0.0001***	0.0001***	
		(0.00001)	(0.00001)	
# of seminars				0.015***
				(0.004)
# of conferences				0.005
				(0.008)
Constant	−1.714	0.980	−0.547	3.475***
	(1.283)	(1.246)	(1.285)	(0.087)
Journal-fixed effects	*No*	*No*	*No*	*Yes*
Publication year-fixed effects	*No*	*No*	*No*	*Yes*
N	4,406	4,406	4,406	4,406
R^2	0.078	0.085	0.090	
Log Likelihood				-22,256.980

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Column 1 through 3 replicate models 1 through 3 of Laband and Tollison (2000, Table 4). Column 3 replicates Panel B of Brown (2005, Table 8), with a slightly different variable definition and without the SSRN control variable. Reported coefficients in column 4 are marginal effects and show the per cent increase in the citation count in response to a 1 unit increase in the independent variable, holding all variables at their mean and setting binary variables to 0. *Auth. total. citations* is the sum of individual citation stocks (according to Scopus) for all authors, measured in the year before publication. *# of authors* and *# of pages* is the simple count of authors and pages, respectively. *Com. total. citations* is the sum of individual citation stocks (according to Scopus) for all commenters, measured in the year before publication. *# of commenter*, *# of seminars* and *# of conferences* is the count of commenters, seminars and conferences acknowledged in the articles' acknowledgment section.

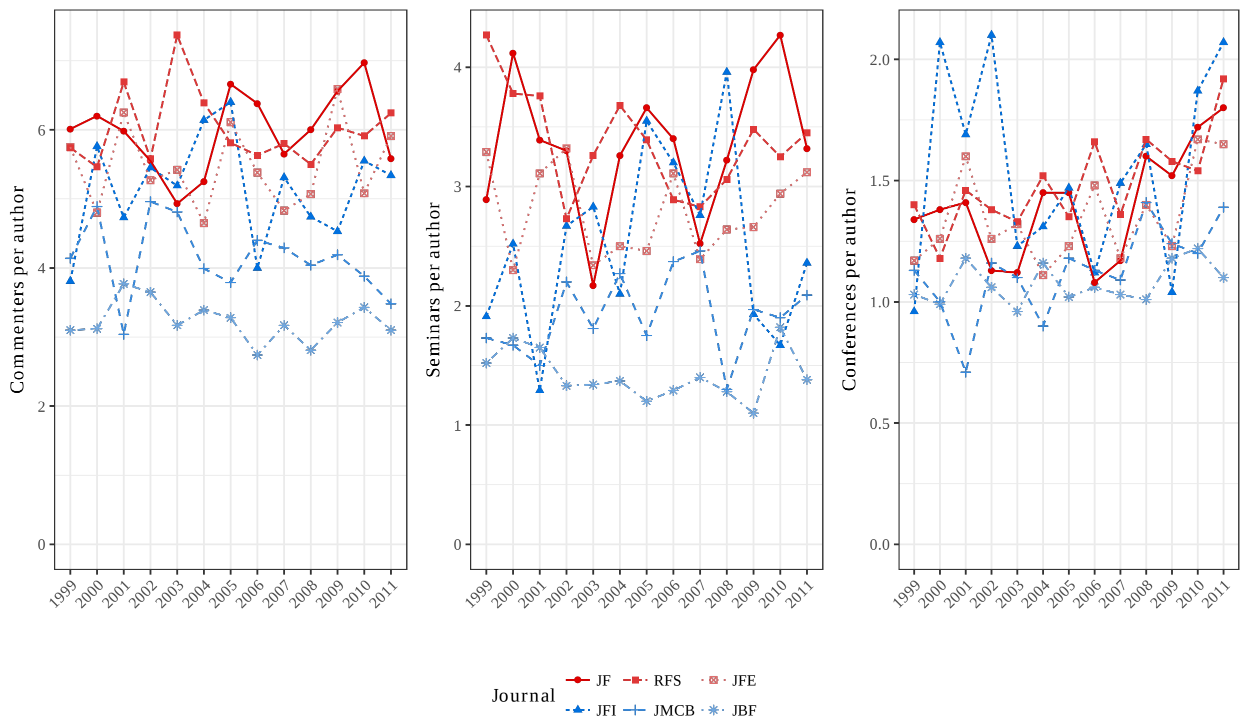
A.4 Figures

Figure 1: Share of articles with and without acknowledgements per journal and year.



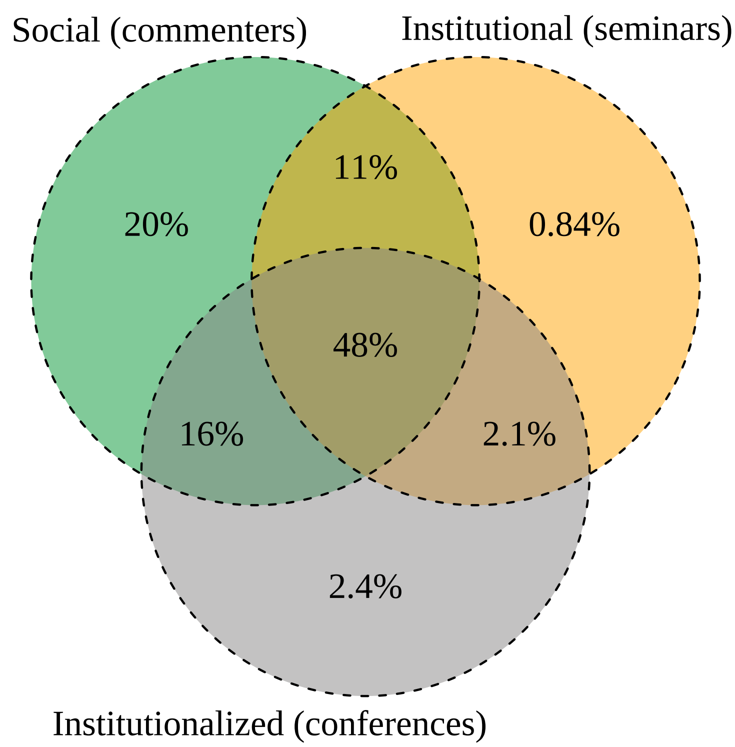
Notes: Graph shows share of articles with (left bar) and without acknowledgments (right bar) for each year. Colors correspond to journals, where red-ish colors refer to the three top journals (JF, JFE, RFS) and blue-ish colors refer to the three field journals (JFI, JMCB, JBF).

Figure 2: Mean number of commenters, seminars and conferences per author over time per journal and year.



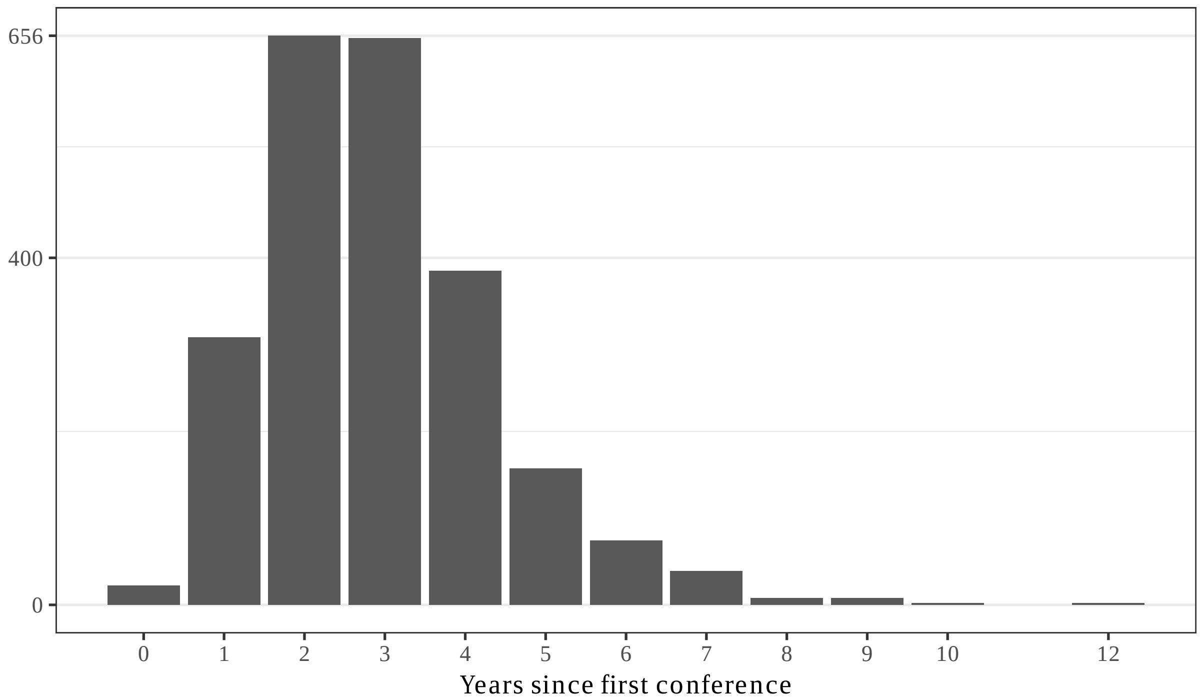
Notes: Graph shows mean number of acknowledged commenters (left plot), seminars (center plot) and conferences (right plot) per journal over time, divided by the number of authors. Colors correspond to journals, where reddish colors refer to top journals.

Figure 3: Share of articles reporting any form of informal collaboration.



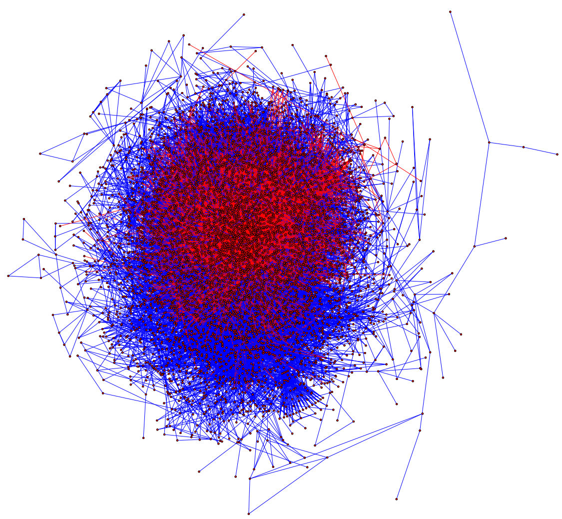
Notes: Venn diagram shows share of articles that report social informal collaboration (colleagues), institutional informal collaboration (seminars) and institutionalized informal collaboration (conferences).

Figure 4: Distribution of estimated age of all research articles, full sample.



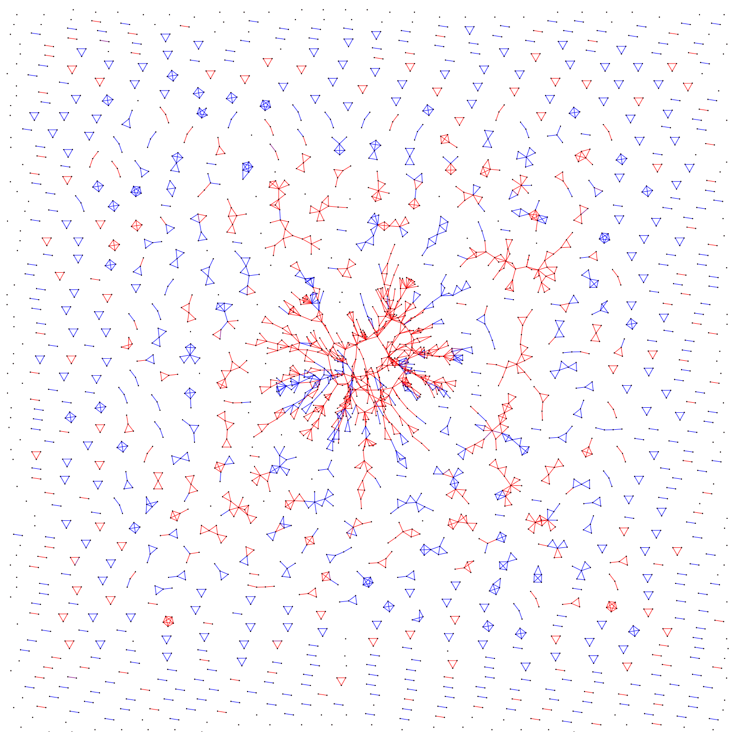
Notes: Histogram shows distribution of estimated age of a article. Age is the difference between the publication year and the earliest year denoted in the names of all acknowledged conference. We only considered research articles, that list at least one conference including the year in the acknowledgment section.

Figure 5: Social network of intellectual collaboration, 2009-2011.



Notes: A link is drawn between an acknowledged commenter and every author of a published research article. Red links indicate that the research article was published in an A journal, while blue indicates a B journal publication. If a link occurs in both an A journal and a B journal, which is a rare event, it is colored in purple. Only the giant component is shown. Graph representation using an adapted Fruchterman-Reingold algorithm.

Figure 6: Social network of co-authorship, 2009-2011.



Notes: A link is drawn between every authors of a published research article. Red links indicate that the research article was published in an A journal, while blue indicates a B journal publication. If a link occurs in both an A journal and a B journal, which is a rare event, it is colored in purple. Graph representation using an adapted Fruchterman-Reingold algorithm.

References

Adams, J. D. (1990), 'Fundamental Stocks of Knowledge and Productivity Growth', *Journal of Political Economy* **98**(4), 673–702.

Azoulay, P., Graff Zivin, J. S. and Wang, J. (2010), 'Superstar Extinction', *The Quarterly Journal of Economics* **125**(2), 549–589.

Ballester, C., Calvó-Armengol, A. and Zenou, Y. (2006), 'Who's Who in Networks. Wanted: The Key Player', *Econometrica* **74**(5), 1403–1417.

Bonacich, P. (1987), 'Power and Centrality: A Family of Measures', *American Journal of Sociology* **92**(5), 1170–1182.

Borokhovich, K. A., Bricker, R. J. and Simkins, B. J. (2000), 'An Analysis of Finance Journal Impact Factors', *The Journal of Finance* **55**(3), 1457–1469.

Bramoullé, Y. and Kranton, R. (2007), 'Public goods in networks', *Journal of Economic Theory* **135**(1), 478–494.

Bramoullé, Y., Kranton, R. and Amours, M. D. (2014), 'Strategic Interaction and Networks', *The American Economic Review* **104**(May), 898–930.

Brown, L. D. (2005), 'The importance of circulating and presenting manuscripts: Evidence from the accounting literature', *Accounting Review* **80**(1), 55–83.

Card, D. and DellaVigna, S. (2013), 'Nine Facts about Top Journals in Economics', *Journal of Economic Literature* **51**(1), 144–161.

Cohen, W. M. and Levinthal, D. A. (1989), 'Innovation and Learning: The Two Faces of R & D', *The Economic Journal* **99**(397), 569–596.

Cohen, W. M., Nelson, R. R. and Walsh, J. P. (2002), 'Links and Impacts: The Influence of Public Research on Industrial R&D', *Management Science* **48**(1), 1–23.

Colander, D. (1989), 'Research on the Economics Profession', *Journal of Economic Perspectives* **3**(4), 137–148.

Cronin, B. (1995), *The Scholar's Courtesy: The Role of Acknowledgment in the Primary Communication Process*, Taylor Graham, Los Angeles, CA.

Debreu, G. and Herstein, I. N. (1953), 'Nonnegative Square Matrices', *Econometrica* **21**(4), 597–607.

Ductor, L. (2015), 'Does Co-authorship Lead to Higher Academic Productivity?', *Oxford Bulletin of Economics and Statistics* **77**(3), 385–407.

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

Ellison, G. (2002), 'The Slowdown of the Economics Publishing Process', *Journal of Political Economy* **110**(5), 947–993.

Freeman, L. C. (1978), 'Centrality in social networks conceptual clarification', *Social Networks* **1**(3), 215–239.

Georg, C.-P. and Rose, M. E. (2015), Mirror, Mirror, on the Wall, Who Is the Most Central of Them All?, mimeo.
URL: *http://www.ssrn.com/abstract=2709107*

Girotra, K., Terwiesch, C. and Ulrich, K. T. (2010), 'Idea Generation and the Quality of the Best Idea', *Management Science* **56**(4), 591–605.

Goyal, S. and Moraga-González, J. L. (2001), 'R&D Networks', *The RAND Journal of Economics* **32**(4), 686–707.

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

Graham, B. S. (2015), 'Methods of Identification in Social Networks', *Annual Review of Economics* **7**(1), 465–485.

Green, R., O'Hara, M. and Schwert, G. W. (2002), 'Joint Editorial: Advice for Authors', *The Journal of Finance* **57**(2), 1031–1032.

Haeussler, C. and Sauermann, H. (2016), 'The Division of Labor in Teams: A Conceptual Framework and Application to Collaborations in Science', *NBER, Working Paper*.

Hamermesh, D. S. (1992), 'The Young Economist's Guide to Professional Etiquette', *Journal of Economic Perspectives* **6**(1), 169–179.

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

Hudson, J. (1996), 'Trends in Multi-Authored Papers in Economics', *Journal of Economic Perspectives* **10**(3), 153–158.

Jackson, M. O. (2014), 'Networks in the Understanding of Economic Behaviors', *Journal of Economic Perspectives* **28**(4), 3–22.

Jackson, M. O. and Zenou, Y. (2015), Games on Networks, pp. 95–163.

Jaffe, A. B. (1989), 'Real Effects of Academic Research', *The American Economic Review* **79**(5), 957–970.

Jones, B. F. (2009), 'The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?', *The Review of Economic Studies* **76**, 283–317.

König, M. D., Liu, X. and Zimmermann, C. (2015), 'Coauthorship Networks and Research Output', *University of Zürich, Working Paper Series*.

König, M., Liu, X. and Zenou, Y. (2014), 'R&D Networks: Theory, Empirics and Policy Implications', *CEPR, Discussion Paper*.

Laband, D. N. and Tollison, R. D. (2000), 'Intellectual Collaboration', *Journal of Political Economy* **108**(3), 632–661.

Levin, S. G. and Stephan, P. E. (1991), 'Research Productivity Over the Life Cycle: Evidence for Academic Scientists', *The American Economic Review* **81**(1), 114–132.

Mansfield, E. (1991), 'Academic research and industrial innovation', *Research Policy* **20**(1), 1–12.

Manski, C. F. (1993), 'Identification of Endogenous Social Effects: The Reflection Problem', *The Review of Economic Studies* **60**(60), 531–542.

Mullahy, J. (1986), 'Specification and testing of some modified count data models', *Journal of Econometrics* **33**(3), 341–365.

Oettl, A. (2012), 'Reconceptualizing Stars: Scientist Helpfulness and Peer Performance', *Management Science* **58**(6), 1122–1140.

Perry, M. and Reny, P. J. (2016), 'How To Count Citations If You Must', *The American Economic Review* **106**(9), 2722–2741.

Phelps, C., Heidl, R. and Wadhwa, A. (2012), 'Knowledge, Networks, and Knowledge Networks: A Review and Research Agenda', *Journal of Management* **38**(4), 1115–1166.

Powell, W. W. and Grodal, S. (2005), Networks of Innovators, *in* J. Fagerberg, D. C. Mowery and R. R. Nelson, eds, 'The Oxford Handbook of Innovators', chapter 3, pp. 153–158.

Singh, J. and Fleming, L. (2010), 'Lone Inventors as Sources of Breakthroughs: Myth or Reality?', *Management Science* **56**(1), 41–56.

Waldinger, F. (2012), 'Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany', *The Review of Economic Studies* **97**(2), 838–861.

Wang, B.-Y. and Xi, B.-y. (1997), 'Some InequallUes for Singular Values of Matrix Products*', *Linear Algebra and its Applications* **264**, 109–115.

Wuchty, S., Jones, B. F. and Uzzi, B. (2007), 'The Increasing Dominance of Teams in Production of Knowledge', *Science* **316**(5827), 1036–1039.