Georg, Co-Pierre; Opolot, Daniel; Rose, Michael

Conference Paper

Discussants

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2019: 30 Jahre Mauerfall - Demokratie und Marktwirtschaft - Session: Industrial Organisation II, No. B25-V1

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Discussants *

Co-Pierre Georg[1,2], Daniel C. Opolot[1], and Michael E. Rose[†3]

[1]University of Cape Town

[2]Deutsche Bundesbank

[3]Max Planck Institute for Innovation and Competition

This study is the first to examine the role of discussants in academic knowledge production. Comparing articles of similar quality with and without discussants, we find that having a discussant increases a paper's probability of publication in prestigious journals, but not its citation count. Conditional on having a discussant, citation count and probability of publication in a prestigious journal increase in the discussants' prolificness. This supports the existence of a *quality* channel through which discussants improve the inherent quality of a paper. Conversely, we do not find evidence for the existence of a *diffusion* channel whereby papers garner more citations because discussants diffuse information about the paper within their social network.

Keywords: Discussants, Intellectual collaboration, Academic Impact, NBER Summer Institutes

JEL Classification: A14, D83, G00

*We wish to thank Pierre Azoulay, Jeff Furman, Fabian Gaessler, Dietmar Harhoff, Adam Jaffe, Danielle Li and Atif Mian, seminar participants at Maastricht University and University of Cape Town as well as conference participants at the 2016 Annual Meeting of the European Economic Association, the 2016 Africa Meeting of the Econometric Society and the INET-YSI "Innovation, Economic Complexity and Economic Geography" Workshop at MIT for helpful comments and feedback. The views expressed in this paper do not necessarily reflect the views of Deutsche Bundesbank.

†Corresponding Author. E-Mail: michael.rose@ip.mpg.de

1 Introduction

Informal intellectual collaboration—receiving feedback and commentary from colleagues on ongoing research papers—is commonplace in economic research.[1] Early empirical work shows that more informal intellectual collaboration, measured as the number of commenters, seminars, and conferences, is associated with a higher likelihood that a paper is accepted for publication and a greater number of citations once it is published (Laband and Tollison, 2000; Brown, 2005). The direction of causality and the channel through which citations increase are not well understood, however.

This paper uses the assignment of discussants at top financial economics conferences, the National Bureau for Economic Research's Summer Institute (NBER SI), to understand their role in academic knowledge production. Our results suggest the existence of a *quality channel* through which dedicated discussants improve the quality of a paper. Specifically, (i) we find a positive effect on a paper's academic success of having a discussant (and possibly a discussion with the audience) versus a discussion with the audience only; (ii) we find a similar positive effect of having a more prolific discussant; and (iii) test the existence of an additional *diffusion channel* through which discussants disseminate information about the paper to their colleagues. Thus, while we find evidence for the existence of the quality channel, we do not find evidence for the existence of a diffusion channel.

We use two related strategies to first identify the effect of having a discussant and then to identify the effect of discussant characteristics. We argue that the effect of having a discussant is well identified if a) authors do not base the decision which Summer Institute to apply for primarily on the fact whether the SI will feature discussants, i.e. there is no sorting by the authors to SIs; and b) if the quality distribution of papers accepted in Summer Institutes with discussants is the

[1] We use the terms *"informal intellectual collaboration"* and *"informal collaboration"* interchangably throughout the paper.

same as in Summer Institutes without discussants, i.e. there is no sorting by the SI organizers of papers into sessions with and without discussants. We argue that both conditions hold in our setup, given the highly competitive and prestigious nature of the Summer Institutes. If there are strategic considerations beyond seeking a good topical fit, it is more reasonable for authors to apply for Summer Institutes that editors of their target journals attend. It is also reasonable to assume that papers presented at NBER SIs represent the complete upper tail of the quality distribution of the universe of current research papers.

Similarly, we argue that the effect of discussant characteristics is well identified if a) the organizer match a discussant to papers and authors based primarily on topical fit (i.e. whether the discussant knows the relevant literature, methods, data, etc.) rather than discussant characteristics; and b) if discussants discuss a paper irrespective of the session it is in. Or, in other words, if the covariance between discussant and author characteristics is zero. We provide evidence that such assortative matching is absent in our data, at least for observable discussant characteristics at the time of the Summer Institute. We rule out any form of sorting by discussants which would arise if potential discussants decline to discuss a paper.[2] This is because being a discussant at the NBER is also considered very prestigious and a signal of high standing within the profession. However, a confounding factor arises from the ability of the organizers to attract discussants. We introduce fixed effects for the corresponding NBER working group.[3] Additionally, we cluster standard errors around the NBER working group to account for unobserved heterogeneity, e.g. in the form of the organizer's professional network (Abadie et al., 2017).

Our main data source are papers presented at the NBER Summer Institutes and subsequently published. NBER Summer Institutes are small workshops with an extremely low acceptance rate, and the accepted papers are therefore of relatively homogeneous quality. For this group of papers, we extract paper-specific details such as the authors' identity, whether a paper had

[2] Indeed organizers have confirmed that discussants rarely reject an invitation.

[3] Organizers seldom change during our sample period, implying that group-fixed effects capture organizer effects as well.

a discussant and if so, the identity of the discussant. We then conduct a follow-up search to collect additional information regarding the characteristics of discussants (e.g. productivity as measured by their Euclidean index of citations, experience). To construct information on the extent to which discussants engage in informal collaboration, we use data from the title pages of 5,759 full research articles from six scholarly journals in financial economics published between 1997 and 2011, initially collected by Georg and Rose (2018).

We use three dependent variables as measures of success of an academic paper: The *total citation count* until January 2018 according to Scopus, whether the publishing journal is one of the commonly denoted "top" journals in financial economics, and the *Journal Impact Factor*.[4] The latter two measures relate directly to the journal as a measure of quality. The journal a paper is published in is arguably the most important metric to evaluate academic economists, for example in tenure decisions.[5]

We find that having a discussant and a shorter discussion with the audience, versus only having a discussion with the audience, increases the chances of publishing in better journals, but does not increase the citation count for published papers. Conditional on having a discussant, citation count and top journal publication probability increase in the discussant's productivity. The effect is statistically and economically significant: A one standard deviation increase in the discussant's prolificness[6] is associated with a 26% increase in citation count, corresponding to 28 more citations. The probability of publishing in a top journal similarly increases by over 25 percentage points. The quality of a journal a paper gets published in (measured as the journal's impact factor) does not increase in discussant characteristics.

[4] In appendix B we perform additional regressions using three different measures of academic success for a smaller sample. These measures are citations, downloads and article views on RePEc, which allows to study the effect right after the publication. The sample includes 238 papers that have been pre-published as NBER working paper.

[5] Journal impact factor as a metric are not without criticism, though, as for example the discussion in Heckman and Moktan (2018) shows.

[6] Our measure of prolificness is the Euclidean index of citations as proposed by (Perry and Reny, 2016). It is a function of citations up until a given year of individual citations to all papers published until that year.

There are two possible channels how a discussant could increase a paper's citation count: a *quality* channel and a *diffusion* channel. The quality channel is straightforward: receiving feedback and commentary from colleagues improves the quality of a paper, and hence its likelihood of getting cited and published in good journals. That is, we expect that highly productive and experienced colleagues possess broad academic knowledge and skills, as well as expertise on how to structure papers for top journal publication. Receiving feedback and commentary from well-published colleagues should thus significantly improve the academic success of a paper. According to the diffusion channel, an observed increase in citation count could result simply because discussants tell their colleagues about the paper.[7] Hence, more widely diffused papers have a higher chance of getting cited and ultimately being published in good journals. Since a paper's commenters are the first to learn about the paper and its quality (because they see an early draft of the paper) their ability to disseminate information about the paper could hence be crucial.

To examine the diffusion channel, we develop a simple and direct measure–*neighborhood centrality*–of an individual's capacity to diffuse information to the wider network. For each discussant in the network, the neighborhood centrality counts the number of all nodes within a given distance from a researcher, while discounting distant connections.[8]

We use two social networks to compute a discussant's neighborhood centrality: A co-author network derived from published papers in all of Economics, and a network of informal collaboration derived from acknowledgements on published papers in Financial Economics (Georg and Rose, 2018). In practice, information dissemination occurs not only via formal ties in the

[7] This form of network diffusion is akin to the diffusion of information about products, practices and services (Richardson and Domingos, 2002; Duflo and Saez, 2003; Bandiera and Rasul, 2006; Conley and Udry, 2010; Banerjee et al., 2013; Cai et al., 2015; Beaman et al., 2018). These papers show that information sharing in networks of friendship, peer-peer, colleagues and family relations, influences individual decisions, and hence, successful diffusion of products, services and practices. Following a similar line of argument, we seek to test whether information sharing, initiated by discussants, in networks of intellectual collaboration influences the success of a paper.

[8] As we will discuss later, several other centrality measures developed in the literature do not apply to the form of diffusion we wish to examine.

form of co-authorship, but also via less formal, but more common ties, such as commenting on a colleagues' work. To explicitly account for this possibility, we include the network of informal collaboration within the scientific community we focus on: Financial Economists. Even when accounting for informal intellectual collaboration as a possible diffusion channel, we find no evidence to support the existence of such a channel.

We study a special form of informal collaboration with academics, namely commenters that are assigned to the authors. The assignment of discussants at NBER SIs not only provides us with a treatment effect, it also helps us to overcome the inherent endogeneity in the formation of academic collaboration networks because authors do not typically choose who their discussant is (Graham, 2015). Despite an exponential growth of the extent of informal collaboration in economics,[9] few papers have examined its implications for knowledge production and on the level of individual papers. Waldinger (2010) documents a fall in output of doctoral students who were left behind when superstar German scientists emigrated during the Nazi era. In a follow-up paper, Waldinger (2012) finds no effect on the output of fellow professors left behind when their colleagues emigrated. Azoulay et al. (2010) report decreased productivity of co-authors of superstar scientists in biological labs after the superstar dies. In a similar setting, Oettl (2012) finds that losing a co-author with a high degree of helpfulness, i.e. how often one is acknowledged by others, leads to a drop in the quality of a researchers output. While the aforementioned studies estimated the impact of (estimated) informal collaboration on others productivity, few studies considered a single research project as unit of observation. Focusing on Economics and Biology, Laband and Tollison (2000) find that a higher number of commenters is associated with a higher citation count over seven years. Brown (2005) includes other forms of informal collaboration, such as seminar presentations and 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.

[9] Laband and Tollison (2000) document the increase in informal collaboration in economics using featured articles of the Review of Economics and Statistics.

What distinguishes our paper from the literature is that we put informal collaboration under the microscope and use discussants and their properties to identify a) their overall causal effect on the academic paper (as opposed to researchers) and b) distinguish the channels through which informal collaboration impacts the success of a paper. This has immediate implications for the design of academic conferences, but also for the distribution of academic credit and merit. For example, it adds to the notion that not just authors write a paper, but a wider research group to which discussants belong (Ponomariov and Boardman, 2016; Borjas and Doran, 2015).

2 Variable construction

2.1 NBER Summer Institutes

We study the effects of informal collaboration on a paper's academic success using papers presented in finance-related National Bureau of Economic Research (NBER) Summer Institutes. NBER Summer Institutes are annual meetings with presentations of up to 17 manuscripts grouped by NBER working group. Some of the Summer Institutes features discussions by external discussants. The program of every NBER Summer Institute since 2000 is available online and includes the title of each paper and the names of authors and discussants.[10] We focus on the Summer Institutes between 2000 and 2010 of the following NBER working groups: Asset Marketing/Real Estate (AMRE), Asset Pricing (AP), Corporate Finance (CF), Impulse and Propagation Mechanisms (EFCE), Capital Markets in the Economy (EFEL), Forecasting & Empirical Methods in Macro & Finance (EFWW), Household Finance (HF), International Finance & Macroeconomics (IFM), Monetary Economics (ME), Finance & Macro (MEFM), Economics of Real Estate & Local Public Finance (PERE) and Risk of Financial Institutions (RISK).

In total 1,016 presentations took place at a total of 132 Summer Institutes between 2000 and

[10] See http://www.nber.org/summer-institute/.

2010, which include at least 31 multiple presentations of the same paper. However, not every presentation eventually resulted in a publication.[11] Overall, 628 (69%) of the presentations resulted in publication in 84 different journals until January 2019.[12] The most important outlets for the published papers are The Journal of Finance (66 papers) and The American Economic Review (64). They are followed by The Review of Financial Studies (52), Journal of Financial Economics (48) and the Journal of Monetary Economics (41).

Table 1 gives an overview of the number of presentations, by year and NBER group. The share of papers published is highest for the NBER group Asset Pricing: 85% of its presentations resulted in a publication until January 2019.

TABLE 1 ABOUT HERE

From the 628 presentations that were eventually published in a Scopus-indexed journal, we remove 34 papers that were presented more than once. We also remove 13 that were published in other than peer-reviewed journals. We arrive at a final sample size of 615 papers. Of these, 378 (61%) have been discussed by 250 distinct discussants.

2.2 Paper characteristics

For each of the 615 papers, we compute three measures of academic success. The first one is the total count of citations a paper garnered until January 2019, as provided by Scopus (Rose and Kitchin, 2019). Second, we use a binary variable equal to 1 if the paper was published in one of the so-called top three Finance journals or top five Economics journals and 0 otherwise.[13]

[11] We restrict the analysis to publications published in sources indexed by Scopus. Due to their editorial policy, most books and collections are exempt from the database.

[12] Some papers change title. We, therefore, conducted an internet search for each paper based on the authors and abstracts to identify those papers with a different title. We exclude 3 presentations from the analysis because the program does not mention their title.

[13] These journals are: Journal of Financial Economics, Review of Financial Studies, and The Journal of Finance, as well as Econometrica, Journal of Political Economy, The American Economic Review, Quarterly Journal of Eco-

Finally, as a continuous measure of journal quality, we use the Journal Impact Factor of the journal the paper was published in in the year before publication, provided by SCImago.[14] This impact factor indicates for every year the average number of weighted citations (as recorded by Scopus) the journal's publications of the past three years received.[15] The weights correspond to the Journal Impact Factor of the journal of citing article. Thus the algorithm is iterative and uses the Eigenscore method in order to take into account the origin of citing articles.

While being a widespread measure of publication quality or scholarly productivity, its use is not without problems. One of the reasons is that few people decide on acceptance for publication in a journal, which is not without inefficiencies.[16] The other measure, citation count, can be thought of as a market-based measure, where the entire community judges the relevance and quality of a paper whenever it is cited (or not). Indeed, papers published in lower ranked journals regularly garner more citations than papers published in top journals (Oswald, 2010). However, there are other problems, too. Paper citations depend on order effects (Coupé et al., 2010; Feenberg et al., 2017) and already highly cited articles attract more citations, too (Merton, 1968). Since none of the measures is perfect, but all are correlated with true paper quality and relevant for the academic career, we include all three measures.[17]

For each article we also count the number of years between the presentation at an NBER Summer Institute and the final publication. This serves as an–albeit imperfect–indicator to how much the article changed subsequent to the presentation.

nomic and Review of Economic Studies.

[14] Obtained from http://www.scimagojr.com/journalrank.php.

[15] As alternative measures of journal quality, we use the journal's h-index and the average citation count over the previous five years. Results are qualitatively the same.

[16] Gans and Shepherd (1994) for example document publication idiosyncrasies of influential papers.

[17] In appendix B we perform additional regressions using three measures of academic success for a smaller sample. These measures are citations, downloads and article views on RePEc, which allows to study the effect right after the publication. The sample includes 238 papers that have been pre-published as NBER working paper.

2.3 Author and Discussant Characteristics

For each of the 250 discussants and the 943 authors of the papers in our dataset, we compute three variables. We denote those as author and discussant characteristics, respectively. Author and discussant characteristics differ only in the period they are computed for.

The first variable is the Euclidean index of citations and serves as a measure of academic prolificness (Perry and Reny, 2016). The Euclidean index of citations is computed as follows. For each year t, count the total number of citations to each of researcher i's n publications published until and including t, then take the square root of the sum of the squared citation counts. That is, if $c_{k,t}$ is the total citation count for paper k until period t, then the Euclidean index $\text{Euclid}_{i,t}$ of author i at period t is:

$$\text{Euclid}_{i,t} = \sqrt{\sum_{k=1}^{n} c_{k,t}^2} \tag{1}$$

A researcher's Euclidean index hence increases monotonically in the number of publications with positive citation count.[18] As an example, consider an author with a stock of two publications in t, one which received 5 citations until t and the other one 50. The Euclidean index of citations equals $\sqrt{5^2+50^2} \approx 50.25$. If the first paper's citation stock increases from 5 to 10 in $t+1$, and the author publishes another paper that garners 2 citations, the Euclidean index of citations would increase to $\sqrt{10^2+50^2+2^2} \approx 51.03$ in $t+1$.

The second measure we use is academic experience, which for any year t is the number of years since the academic's first publication. As third measure, we use the squared experience in order to capture academic life-cycle effects. All these measures were computed using data from Scopus (Rose and Kitchin, 2019).

[18] Alternative measures would include 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.

2.4 Information Diffusion in Social Networks

We estimate a discussant's ability to diffuse information contained in a paper to their respective network using a two different network measures computed in two different social networks. One network, commonly used in the literature, is the network of co-author relationships. While social networks matter for the diffusion of information, it is not just formal ties that matter. Economists collaborate informally by commenting on each others' work, as Laband and Tollison (2000, 2003); Georg and Rose (2018) argue. Informal ties arise, for example, when one researcher provides commentary on another researcher's paper, or is a discussant at another conference. By including the social network of informal collaboration, we account for a much broader range of possible information diffusion.

The co-author network is inferred from a total of 409 scholarly journals indexed in Scopus. We define the set of journals according to field-wise rankings in Combes and Linnemer (2010). We include every journal that is ranked C or better in any field-wise ranking. From each set of published articles in every year, we include what Scopus labels as research article, conference proceeding or review.

To construct networks of informal collaboration, we use the acknowledgement dataset of Georg and Rose (2018). The dataset contains manually collected acknowledgements to individual researchers from 5,759 research papers published between 1997 and 2011 in six Financial Economics journals with similar topical focus: 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). Informal collaborators include all persons that are acknowledged for help with concept-related tasks.[19] From the list of acknowledged commenters, referees, if thanked by name, and

[19] Cronin (1995) characterizes concept-related tasks as provision of ideas, feedback and commentary. The words that authors use to indicate input on conceptual issues are usually: comments, insights, encouragements, discussions.

the managing editor(s) of the journal for the year of publication and the previous two years were removed. This is to avoid an overestimation of the editors' position in the network.[20] Georg and Rose (2018) furthermore argue that strategic acknowledging is not the driving motive of acknowledging, as empirical observations speak against this hypothesis.[21]

To construct a network of informal collaboration, we connect two researchers i and j with a link $h_{ij} > 0$ whenever one acknowledges the other as a commenter on a paper in the acknowledgement dataset. Links between these authors and commenters are undirected. This is to capture the fact that the author tells the commenter about new but yet unpublished results, or emerging ideas and concepts to build her own research on. Conversely, spillovers from the commenter to the author occur 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. We weigh links to reflect the level of interaction between collaborators, and to account for possible mis-measurement, with the inverse of the number of authors.

Taken together, for each year t we construct the co-author network and the network of informal collaboration using the publications in that year, the previous year $t-1$, as well as in the following year, $t+1$ (this accounts for the publication lag, as the project must have begun earlier). We denote the co-author network by G and the network of informal collaboration by H. Formally, let A_t be the set of papers published in years $\{t+1, t, t-1\}$. To each paper $a \in \{A_t\}$, there is a non-empty set of authors κ_a and a set of commenters ι_a. For each pair $(k,i) \forall k, i \in \kappa_a$, g_{ki} increases by 1. Thus links in the co-author network are weighted by the frequency with which authors jointly publish papers. For each pair $(k,j) \forall k \in \kappa_a, j \in \iota_a$, h_{kj} increases by $1/|\kappa_a|$. If, for example, a commenter has been acknowledged on a paper written by two authors, there would

[20] The vast majority of papers 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 papers a journal publishes, the higher is its editor's observed centrality in the uncorrected data.

[21] According to Hamermesh (1992), strategic acknowledging is an author's attempt to influence an editor in her choice of referees. The underlying belief is that an editor would (not) pick one of the acknowledged commenters are referee.

be two links with weight 1/2 connecting each of the authors with the commenter. If one of the authors acknowledges this commenter on another single-authored paper, the weight of the two links connecting that author and the commenter increases to 3/2. The resulting adjacency matrices G and H are thus symmetric, with the elements on the diagonal equal to 0.

For every $t = 2000, 2001, \ldots, 2010$, we construct both the co-author network and the network of informal collaboration. The 2000 co-author network connects 23,443 researchers and was inferred from 28,705 papers published in 1999, 2000 or 2001. It grows tremendously until 2010: The 2010 co-author network was inferred from 47,401 papers and connects 47,297 researchers. We generate 2000 network of informal collaboration from 891 papers published in either 1999, 2000 or 2001 and consists of 3,534 distinct researchers. This compares to the 2010 network of informal collaboration, which connects 7,475 researchers that have collaborated on 1,889 papers in our dataset.

Being equipped with both networks, we turn to the definition and computation of network measures that suitably capture a discussant's ability to impact the success of the paper through the diffusion channel. The diffusion starts from the discussant, who is among the first people to learn about the existence and quality of the paper. The discussant then tells her immediate colleagues (co-authors and informal collaborators) about the existence and quality of the paper, who then tell their colleagues, and so on. As information about the paper diffuses through the network, researchers who learn about it may decide to cite the paper. To test for the existence of such diffusion effects, we first define network-based measures of a discussant's ability to diffuse information about the paper. That is, a discussant's capacity, due to her position in the network of intellectual collaboration, to diffuse the paper directly and indirectly to the wider academic community.

Although there are several individual centrality measures developed in the literature (e.g. degree, eigenvector, Bonacich, diffusion and betweenness centralities) most of them do not ap-

ply to the type of diffusion we examine. For example, Banerjee et al. (2013) find that diffusion centrality outperforms other measures of centrality in predicting microfinance participation resulting from information sharing. This suggests that the relevant centrality measure depends on the form of diffusion examined. Neighborhood and degree centralities are the most suitable centrality measure for the diffusion mechanism we study. That is, the diffusion process that starts from a given node (a discussant) and propagates outwards (to a discussant's colleagues, and to a discussant's colleagues' colleagues, and so on) with potentially diminishing effects.[22]

We are interested in an information-driven diffusion channel where the sheer knowledge of a paper's existence and quality increases the likelihood to cite it. But this diffusion of knowledge about a paper is different from strategic influencing neighboring nodes' decisions to cite a paper, which is why we cannot use Bonacich centrality directly.

The form of network diffusion we examine germinates from one node (i.e. a discussant) and spreads outwards, with potentially diminishing impact as it spreads further away from the source node. The degree—the number of immediate neighbours a node has—captures only the first step of this process. Our measure–*neighborhood centrality*– captures network-wide diffusion while accounting for potential informational decay and relevance. Specifically, we define neighborhood centrality as a measure of the number of all neighbors of a node within a given radius while discounting for relevance in distant neighborhoods.

Formally, let $k_{i\tau}$ be the number of τ-th order neighbors of i. That is, $k_{i\tau}$ is the number of all nodes at a radius τ from i. Let δ be a discount factor of information decay. Then, the *neighbor-*

[22] Although Bonacich and diffusion centralities also measure an agent's influence with diminishing effects, these measures are consistent with the diffusion of behaviour in strategic settings (Ballester et al., 2006). Eigenvector centrality is consistent with the measure of influence in models of opinion formation through averaging (Golub and Jackson, 2010), while betweenness centrality measures an agent's influence as a "gateway" for information flow across the network.

hood centrality of i is simply defined as:

$$n_i(\delta) = \sum_{\tau=1}^{\infty} \delta^{\tau-1} k_{i\tau}. \tag{2}$$

From (2), the relevance of i's information to her distant neighbors is discounted by a factor δ. One reason for discounting information diffusion across the network is to account for topical and expertise distances. That is, from i's point of view, the first-order neighbors k_{i1} conduct research in topics closely related to i, with potentially very similar methodologies and closely related literature. For neighbors further away, however, the methodologies used and topics investigated diverge, so that distant neighbors of i find her information less relevant. Other reasons for discounting include pure decay in the quality of information as it is passed on from one scientist to the next.

We also see from (2) that the degree is equal to the neighbourhood centrality when $\tau = 1$. Also, the degree is very close to neighbourhood centrality when the discount factor δ is close to zero. This corresponds to the case in which either information about the paper decays rapidly across the network, or there is very little overlap across different research groups. The latter situation occurs, for example, when the scientific community is made up of close-knit but distinct research units with very little interaction between them. This relationship between degree and neighbourhood centrality can thus act as a good indicator of the extent or strength of the diffusion process. Overall, these two measures represent a discussant's ability to diffuse the paper to the wider scientific community.

We compute the two network measures, degree and neighbourhood centrality, from the two networks—co-authorship and informal collaborations networks—defined above. For technical reasons, the neighborhood centrality use only the largest component of the network only.[23]

[23]This is because the computation of the centralities relies on paths and is hence specific to that component. However, centralities across components are not comparable.

The largest component, also called giant component, is the largest group of nodes that are connected via intermediate sequences of nodes and links, called a path.

3 Identifying the Effect of Discussants on Paper Success

We wish to show two causal effects: First, the effect of having a discussant (and potentially a short discussion with the audience) versus having a discussion with the audience only on the academic success of a paper, and second, the effect of discussant prolificness on the academic success. Our two identification strategies are thus related.

Common to both identification strategies is the assumption that the quality of accepted manuscripts is homogeneous. This assumption is difficult to test as the only possible measure of impact, citation count, is usually not observable until publication. However, the high standing of the NBER Summer Institutes within the profession raises the supply of high quality papers to a near-complete level, and the strong selection process strives to select only the very best manuscripts.[24] For this reason, we believe that manuscripts presented at NBER summer institutes represent the complete upper tail of the quality distribution of ongoing research.

Whether a paper is discussed or not at an NBER Summer Institute depends on two consecutive choices: The authors apply to a Summer Institute that features discussants, and the corresponding organizers accept the paper. The effect is identified in the absence of sorting of either of the two decisions makers, where sorting is based on the fact that there are (or are not) discussants. That is, authors must not base the decision which Summer Institute to apply for primarily on the fact whether it will feature discussants and organizers of Summer Institutes without discussants do not accept papers of different quality than organizers of Summer Institutes with

[24] We acknowledge that the process of paper selection, in particular when it comes to papers "on the edge", decisions are complex and can involve political considerations. However, given the universe of all economics research papers, those presented at NBER Summer Institutes are quite homogeneous in quality.

discussants. We argue that first second condition holds because it is more rational for authors to achieve a high topical fit with the Summer Institute's theme. If there are strategic considerations beyond seeking a good topical fit, it is more reasonable for authors to apply for Summer Institutes that editors of their target journals attend. Because of homogenous quality, and the non-sorting of authors at the time of application, the quality distribution of papers at the time of submission should not differ between Summer Institutes with and without discussants.

Which discussant a paper ultimately receives depends on the Summer Institute organizers. Since the assignment is non-random, the effect of discussant characteristics is identified if a) the organizers match discussant to papers and authors based on topical fit rather than characteristics and b) sorting from the side of discussants is absent. Formally, the covariance between discussant traits and author characteristics must be zero. We argue that the second condition holds because being discussant at an NBER Summer Institute is considered very prestigious and thus invited discussants rarely decline.[25]

The first condition requires that the discussant's topic and knowledge (i.e. whether the discussant knows the relevant literature, methods, data, etc.) is more relevant than research strength. This form of assortative matching between discussants and promising papers is probably the biggest threat to our identification strategy. But the identification of promising papers in the manuscript stage and subsequent matching of fitting discussants remains imperfect. To study the existence of assortative matching in our sample, we look for statistically significant relationships between the discussant characteristics and any author or paper characteristics. The discussant characteristics are: Whether there was a discussion, the discussant's Euclidean index of citations, the experience and the degree and the neighborhood centrality, each measured in both the co-author and the network of informal collaboration. Each variable is measured in the year of the discussion (or in the network corresponding to the year of the discussion). Explanatory variables are the manuscript's authors' joint Euclidean index of citations, their joint

[25] Indeed organizers have confirmed that discussants rarely reject an invitation.

experience, their joint squared experience, and the number of authors.[26] Table 2 reports results for of logistic (for the binary variable "has discussion") and negative binomial regressions (for discrete variables), while table 3 reports results of corresponding OLS regressions (for all other variables).

TABLES 2 AND 3 ABOUT HERE

We do not find any evidence for assortative matching when looking at author Euclidean index of citations. The only positive relation we find between author experience and any of the discussant characteristics is with a discussant's degree. The authors' total experience squared is negatively correlated with the paper's probability to have a discussion as well as with the discussant's Euclidean index of citations, although the effect is not clear since the total experience itself is insignificant in both cases. We do find some evidence for assortative matching in the author group size: Papers with more authors are significantly more likely to have a discussion, tend to have younger discussants who are, however, more connected in the network of co-authorship. Our results help alleviate fear of assortative matching, although we cannot fully rule it out, in particular with regards to the size of co-author teams. However, the assortative matching here could be driven by paper-type (theory vs. empirical) as empirical papers tend to have more co-authors than theory papers. Anecdotal evidence suggests that empirical papers can be more intensely contested and debated than theory papers and hence might have a higher propensity to be discussed.

Finally, a confounding factor arises from differting ability of the organizers to attract discussants. Some organizers might, due to their standing or prominence in the field, reach a wider pool of researchers that could discuss a paper. But organizers tend to be highly regarded and well connected academics who know their respective field well beyond their personal network.

[26] We assume the author group size to be constant between the discussion and the publication. We exclude the number of pages here because we do not observe the number of pages prior to publication. Including the number of pages of the published paper does not change the results.

We therefore argue that varying reach within the profession is not a severe issue for our identification strategy. To be on the conservative side, we however introduce fixed effects for the corresponding NBER working group.[27] Additionally, we cluster standard errors around the NBER working group to account for unobserved heterogeneity, e.g. in the form of the organizer's professional network (Abadie et al., 2017).

Discussants are arguably a special form of informal collaborators, and discussants at NBER Summer Institutes might constitute a special form of discussants. The most important difference compared to other forms of informal intellectual collaboration (e.g. feedback from seminar participants and feedback from department colleagues), is that NBER discussants discuss the paper in great detail, conduct own analyses and sometimes spend up to a week preparing the discussion. This characteristic of NBER discussions ensures greater knowledge exchange between discussants and authors. While not every acknowledged commenter puts forward that much effort on others' work, they are, however, representative in a functional sense: Providing feedback and insights. Thus, our results provide a fair estimate for all concept-related informal collaboration.

Finally, we face a selection bias coming from the fact that we only observe published papers. It might be that discussed papers have a lower propensity to be published and those that get published, get published in highly prestigious journals. The discussant would thus act as a gatekeeper or screener. While this constitutes a relevant mechanism, it would bias our results. However, we don't find that having a discussant, not the discussant's characteristics, predict whether the paper will be published or not. These results hold with and without group-fixed effects.

[27] Organizers seldom change during our sample period, implying that group-fixed effects capture organizer effects as well.

4 Empirical Estimation and Results

Our empirical analysis proceeds in three steps. In the NBER discussants sample we identify the effect of informal collaboration with discussants. The discussants characteristics sample is next, which serves to identify the effect of discussant characteristics. We finally assess the impact of discussant network position in the discussant network sample.[28]

4.1 Informal Collaboration with Discussants and Academic Success

To identify the effect of having a dedicated discussant (versus having a general discussion with the audience only) on the academic success of paper i we estimate the following model:

$$\text{Success}_{i,t} = \alpha_0 + \alpha_1 \cdot Paper_i + \alpha_2 Author_{i,t-1} + \beta Discussion_{i,t} + \epsilon_{i,t} \tag{3}$$

where $Success$ is one of our main dependent variables: total Citation Count; whether the paper was published in a Top Journal; and the Journal Impact Factor. $Paper$ is a vector of paper-specific variables that contain the number of authors and the number of pages. $Author$ is a vector controlling for author characteristics and includes Euclidean index of citations (equation (1)), the experience and the square thereof. $Discussion$ is a binary variable equal to 1 if the manuscript was discussed at least once until date t.[29] A positive β would indicate a positive impact of informal collaboration with a discussant on the paper's academic success.

TABLES 4 AND 5 ABOUT HERE

[28] In appendix B we repeat two of the three exercises using three measures of academic success for a smaller sample. These measures are citations, downloads and article views on RePEc. The sample includes 238 papers that have been pre-published as NBER working paper.

[29] Theoretically, a paper could have a discussion *after* it has been published. Due to the long publication cycle in economics and finance (see Ellison (2002)), however, this does not occur in practice.

Table 4 presents all summary statistics for all continuous variables for the sample. Table 5 reports Spearman and Pearson correlations between all variables used in the NBER sample. The average citation count equals 105.4 citations between date of publication and January 2019. 63% of the 615 papers in this sample were published in a top five Economics or top three Finance journal. The average paper in this sample was published in a journal with Journal Impact Factor equal to 8.3.[30] An interesting stylized fact is the age of an article: On average, an article is published in the fourth year after it has been presented at an NBER Summer Institute. This publication lag can go up to as many as thirteen years. Negative experience values for authors mean that all authors have not published before t. Likewise, negative experience values for discussants mean that discussants have not published before or in year z (the year of the discussion). Figure 1 visualizes the experience distribution. About 7% of all discussions were delivered by discussants with negative or zero experience. These are typically very junior discussants, often those just out of their PhD. But also practitioners who follow different publication strategies than academics are in our dataset and can have negative experience according to our measure.

FIGURE 1 ABOUT HERE

Table 6 presents the estimates of three econometric models. In column (1), we estimate a negative binomial model with the paper's citation count until January 2019 as dependent variable. Our variable of interest, whether the article was discussed during an NBER Summer Institute is statistically insignificant. In column (2), we estimate a logistic regression with journal class dummy as dependent variable. Journal class refers to top five Economics or top three Finance journals. Having a discussant statistically significantly increases the likelihood of publishing in a top journal by a factor of two, which is remarkable. Finally, column (3) presents results from

[30] To put this number into perspective: A typical journal with an Impact Factor equal to 8.5 was the Review of Economic Studies in 2008. The highest Journal Impact Factor equals 27.2 and characterizes the Quarterly Journal of Economics in 2014.

an OLS model with the Journal Impact Factor in the year of publication as dependent variable. Again, having a discussant statistically significantly affects the quality of the publishing journal by a magnitude of 1.923. This value is close to the difference in the journal impact factors of Econometrica (19.932) and The Journal of Finance (18.318) in 2017.

TABLE 6 ABOUT HERE

While there is an effect of having a discussant on whether and where a paper gets published, we do not find a significant effect of having a discussant on subsequent citations. This suggests that having a discussant indeed improves the quality of the paper—assuming that the journal publication process indeed serves as a screening mechanism to identify papers of high academic quality—but not necessarily the awareness of the paper within the profession. We have only used whether or not a paper had a discussant, so the next natural question to ask is whether discussant characteristics have an impact on the academic impact (publication success and number of citations) of the paper.

4.2 Prolificness and Academic Success

To understand the impact of a discussant's characteristics on the academic success of a paper, we regress the three paper success outcomes on three discussant characteristics. We focus on 378 papers that had a discussant at an NBER Summer Institute prior to their publication. The econometric specification relates paper success to discussant characteristics in the year of the discussion, controlling for paper and author characteristics:

$$\text{Success}_{i,t} = \alpha_0 + \alpha_1 \cdot Paper_i + \alpha_2 Author_{i,t-1} + \beta Discussant_{i,z\leq t} + \epsilon_{i,t}, \tag{4}$$

where $Success$ is, as before, one of our main dependent variables: Citation Count; whether the

paper was published in a Top Journal; and the Journal Impact Factor. $Paper$ is again a vector of paper-specific variables that contain the number of authors, the number of pages, and the number of discussants. $Author$ is a vector controlling for author characteristics and includes the Euclidean index of citations (equation (1)), the experience and the square thereof. These variables are measured in the year before publication, $t-1$. $Discussant$ is now a vector containing discussant characteristics as measured in the year of the discussion z (which is before the date of the publication t):[31] The Euclidean index of citations of the discussant, the experience of the discussant and the square thereof. In case a paper was discussed more than once we use the joint Euclidean index of citations, as well as the total experience of all discussants.

TABLES 7 AND 8 ABOUT HERE

Summary statistics for this sample are reported in table 7 and correlation coefficients in table 7. Despite the somewhat smaller sample size, the sample has very similar characteristics as the previous sample, with 106.4 citations for the average paper, 63% published in a top Economics or top Finance journal, and an average Journal Impact Factor of 9.3. The average discussant has a Euclidean index of 186.5 and on average 11.4 years of experience in academic publishing. We find low correlation coefficients between author characteristics and the characteristics of discussants, which again indicates the absence of assortative matching.

In table 9, we report results of three regressions corresponding to model (4). Similar to the previous estimation, the first column reports marginal effects of a negative binomial regression with citation count as dependent variable. The second column reports coefficients of a logit regression with the dependent variable being top journal. Finally, column (3) reports OLS coefficient with the Journal Impact Factor as dependent variable. In each specification we cluster standard errors around the NBER group to which the paper belongs.

[31] Theoretically, it is possible that a paper is discussed after it was published, but practically this has never happened in our data.

TABLE 9 ABOUT HERE

While the coefficients for author characteristics change very little across specifications, we find statistically significant effects of discussants' productivity on all academic outcomes. For a one standard deviation increase in the discussants' Euclidean index of citations (260.82) from the mean, the citation count for the average paper increases in expectation by $260.82 \times 0.001 \approx 26\%$ from the mean ($\approx$ 28 citations) until January 2019. Similarly, the probability of publishing in a top journal for the average paper increases by 26 percentage points. The discussant's experience, on the other hand, is not statistically significantly correlated with total citations, the probability of publishing in a top journal, or the journal impact factor a paper gets published in eventually.

4.3 Network Position and Citation Count

So far, our results give credence to the notion that discussants indeed increase a paper's success by improving its quality. However, discussants might also improve a paper's success through their positions in the network of intellectual collaboration. We use degree and neighbourhood centralities (see Section 2.4) to capture a discussant's ability to increase a paper's citation count through the *diffusion channel*. By being the first to learn about the existence and quality of the paper, discussants may not only decide to cite the paper themselves, but also let their immediate colleagues (both formal and informal collaborators) know about it. This can happen through regular discussions, or through an indirect process where a discussant's colleagues learn about the paper because the discussant cites it. The main variable of paper success affected by this channel is the citation count.

We then examine the following two hypotheses in relation to the two channels through which a discussant's position in the network of intellectual collaboration may affect a paper's success:

(i) Discussants are among the first to cite a paper and they cite a significant number of the papers they discuss.

(ii) A discussant's ability to diffuse the paper to the scientific community positively affects the citation count.

To assess the validity of hypothesis (i) above, we use Scopus data to compare the citation patterns by discussants and non-discussants. Figure 2 plots the number of citations from the two categories (i.e. discussants and non-discussants) by time from publication. We observe that discussants tend to cite the paper earlier than non-discussants—but also stop citing an article earlier. Half of the citations from discussants' publications occur after 4 years of the discussed paper's publication, while half of the citations from non-discussant' publications occur after 6 years. In total, discussants usually cite the paper only up to 12 years after publication, while non-discussants start citing the paper three years before publication and end 17 years later.[32] The weighted average time until a citation from a discussant occurs is 4.46 years, while that for non-discussants is 5.88.[33] These descriptive statics are thus consistent with the hypothesis that discussants tend to cite the paper earlier than everyone else, which could support the existence of a diffusion channel.

FIGURE 2 ABOUT HERE

We now examine the second part of hypothesis (i) above. That is, whether the citation count due to discussants is non-negligible compared to the total citation count. For each discussed paper, we count the number of discussants that cite the paper upon publication, and compare it to the total citation count. Out of 378 discussed and subsequently published papers, 88 (20%)

[32] For a paper published in 2001, this is the maximum time range we can observe, since our citation data ends in January 2019.

[33] The measure corresponds to Macaulay duration used to characterize bonds. We define it as $\sum_t^n t \cdot c_t / \sum_t^n c_t$ for n periods before and after publication year in which the publication received at least one citation. c_t is the number of citations in period t.

are cited at least once by its discussants. While this number is non-negligible, it does not tell us whether this is higher or lower than the unconditional probability that a discussant would have cited the paper without having discussed it. But this counterfactual is unobservable.

To test hypothesis (ii), we compute the two network measures (degree and neighborhood centrality) in both the co-author network and the network of informal intellectual collaboration corresponding to the year of the discussion. We model the relationship of a paper's citation count and the discussant's network position, controlling for her prolificness, in the following way:

$$\begin{aligned} \text{Total Citations}_{i,t} &= \alpha_0 + \alpha_1 \cdot Paper_i + \alpha_2 Author_{i,t-1} + \alpha_3 \cdot Discussant_{i,z<t} \\ &\quad + \beta \cdot Network_{i,z<t} + \epsilon_{i,t}. \end{aligned} \tag{5}$$

As before, $Paper$ is a vector of paper-specific variables that contain the number of authors, the number of pages, and the number of discussants. $Author$ is a vector controlling for author characteristics and includes Euclidean index of citations, the experience and the square thereof. These variables are measured in the year before publication, $t-1$. $Discussant$ is a vector containing discussant characteristics measured in the year of the discussion: The Euclidean index of citations, the experience and the square thereof. Finally, we add $Network_{i,z}$ as a vector containing the discussant's network position in the either the co-author network or the network of informal collaboration corresponding to the year of the discussion, z: degree and neighborhood centrality. We can rule out the null hypothesis that there are no diffusion effects, if β has a positive sign and is statistically significantly different from zero.

As outlined in Section 2, we use both the co-authorship network and the network of informal intellectual collaboration to explicitly account for the possibility that information can be diffused via along informal social ties. While we run separate regressions for the co-author network and the network of informal collaboration, we report summary statistics jointly. Tables 10 and 11

report summary statistics and correlation coefficients. Each complete observation is a paper published until January 2019, that was discussed at a finance-related NBER Summer Institute by a discussant who is part of the giant component of the co-author network (or the social network of informal collaboration) corresponding to the year of the discussion.

TABLES 10 AND 11 ABOUT HERE

Table 12 reports results of a negative binomial regression corresponding to model (5) using variables inferred from the co-author network. Reported coefficients are marginal effects evaluated at the sample mean with standard errors clustered around NBER group. Table 13 is the counterpart using network variables for the network of informal collaboration.

TABLE 12 ABOUT HERE

TABLE 13 ABOUT HERE

In both cases, co-author and informal collaboration networks, we find that both degree and neighbourhood centralities are not statistically significant. This means we fail to reject the null hypothesis that there is no diffusion channel. Although there is a potential for the discussants to diffuse the paper to the wider scientific community (in that they are the first to learn about and cite the paper), their contribution in this regards is not statistically significantly different from zero. One possible natural explanation is that other methods through which researchers find papers (e.g. internet search using google scholar, or simply going to the library) are equally or more important than learning about new papers through their social network.

5 Conclusion

Informal collaboration is not just common in economic research, it is ubiquitous. Despite this, very little is known about how informal collaboration impacts the quality and dissemination of research output. A handful of papers that attempt to examine this problem find a positive correlation between the extent of informal collaboration (i.e. the number of acknowledged commenters, seminars and conferences) and citation count. Insightful as these seminal findings are, they do not tell us a lot about the underlying economic mechanisms and the direction of causality.

To identify the causal impact of informal collaboration on research output, we focus on one specific form of informal collaboration: discussions in workshops and conferences, and in particular, discussions at prestigious scholarly conferences in economics, the NBER Summer Institutes. We first show that papers benefit from having a discussant (i.e. who reads and discusses the paper in detail versus having a general discussion with the audience) in terms of an increased probability to be published in prestigious journals. Papers' citation counts increase more when the paper is discussed by a more prolific discussant. This finding lends support to the existence of a quality-improving channel of informal collaboration. On the other hand, we find no evidence on whether increased citation counts are driven by the discussant's ability to diffuse information within the profession.

Our findings contribute to the general understanding of the processes of knowledge production. Understanding these processes is highly relevant as economies shift towards the production of intangible and increasingly knowledge intensive goods.

A Algorithmic steps for computing neighborhood centralities

Neighborhood centralities are computed from a non-weighted adjacency matrix representing a network of informal collaboration. Let G denote the respective adjacency matrix of a network of informal collaboration. G is a zero-one matrix with elements $g_{ij} = g_{ji} = 1$ if a link exists between i and j and zero otherwise. For any arbitrary matrix A, let $\Delta_{\geq 1}(A)$ be a zero-one matrix derived from A by replacing all elements of A that are greater or equal to one with one, and zero otherwise. Let $\Delta_{\leq 0}(A)$ on the other hand be a zero-one matrix derived from A by replacing all negative elements of A with zero. And let I denote the identity matrix.

Let $M_{[\tau]}$ for $\tau = 1, 2, \cdots$ be a modified adjacency matrix representing τ-order neighbor-relations. That is, at each τ, the ijth element of $M_{[\tau]}$ is one if j is at radius τ away from i, and zero otherwise. The steps for computing the sequence $\{M_{[\tau]}\}_\tau$ for $\tau = 1, 2, \cdots$ are as follows.

1. When $\tau = 1$, $M_{[1]} = G$.
2. When $\tau = 2$, let $W_{[2]} = M_{[1]} \times G$; $Z_{[2]} = \Delta_{\geq 1}\left(W_{[2]}\right) - \left(I + M_{[1]}\right)$, then $M_{[2]} = \Delta_{\leq 0}\left(Z_{[2]}\right)$
3. When $\tau = 3$, let $W_{[3]} = M_{[2]} \times G$; $Z_{[3]} = \Delta_{\geq 1}\left(W_{[3]}\right) - \left(M_{[1]} + M_{[2]}\right)$, then $M_{[3]} = \Delta_{\leq 0}\left(Z_{[3]}\right)$

   $\cdots$, $\cdots$ $\cdots$, $\cdots$

t. When $\tau = t$, let $W_{[t]} = M_{[t-1]} \times G$ and $Z_{[t]} = \Delta_{\geq 1}\left(W_{[t]}\right) - \left(M_{[t-2]} + M_{[t-1]}\right)$, then $M_{[t]} = \Delta_{\leq 0}\left(Z_{[t]}\right)$

   $\cdots$, $\cdots$ $\cdots$, $\cdots$

   Continue until all entries of $M_{[\tau]}$ are zero.

Let $\mathbf{e}$ be a column vector of ones, and let T be the value of τ at which all elements of $M_{[\tau]}$ become

zero. Then the vector of neighborhood centralities with discount factor δ is defined as

$$\text{Neighborhood}(\delta) = \sum_{\tau=1}^{T} \delta^{\tau-1} M_{[\tau]} \mathbf{e} \tag{6}$$

B Citations, downloads and article views on RePEc

As an addition to the three measures of academic success, this section presents three other measures relating to NBER Working Papers in our sample. 358 papers out of the 615 papers in our sample previously circulated as NBER Working Paper. Of these, 238 appeared as NBER working paper before final publication. Studying this sample yields insights into the effects of discussants on a paper even before publication.

From the Research Papers in Economics (RePEc) and its associated projects Citations in Economics (CitEc) and Access Statistics for RePEc Services (LogEc), we obtain three additional pieces of information for each of these 238 papers: The total number of citations to the working paper, the total number of downloads of the working paper via RePEc and the total number of views of the database entry on the working paper. Each of these variables were counted until January 2019. The disadvantage is the implied smaller sample size and a different representation of the true academic success.[34]

First we estimate a model similar to (3), where the dependent variable is one of "Total citations (RePEc)", "Total downloads" and "Total views". Paper characteristics includes the number of authors only (since "Age" and "Age2" are not necessary and "# of pages" refers to the number of pages of the published paper). Author characteristics were counted in the year the paper became a NBER working paper. Summary statistics and correlations for this sample are given

[34] An ideal setting would entail observing measures of academic success of the working paper before the year of the discussion. A Diff-in-Diff regression would then compare working papers with and without discussant before and after the presentation at an NBER Summer Institute. There are, alas, only 17 papers that existed as working paper before they have been presented at an NBER Summer Institute. Only 4 of them did not have a discussant.

in tables 14 and 15. The mean number of citations equals 10.5, the mean number of downloads equals 79.7 and the mean number of views equals 213.5. The correlation between these measures is relatively low (0.54 at most).

TABLES 14 AND 15

Table 16 presents the results of negative binomial regressions for each of three dependent variables. Having a discussant has no effect on the number of views, but we find a statistically significant negative effect on the number of RePEc citations and a weakly statistically significant effect on the number of downloads. If the prior is that discussants increase the number of citations, then this finding rejects that hypothesis in a more extreme way than the one presented in section 4.1 (according to which discussants have no effect on post-publication citation counts).

TABLE 16 ABOUT HERE

A similar exercise to the one presented in section 4.2 is to estimate negative binomial regressions similar to model (4). Dependent variables are "Total citations (RePEc)", "Total downloads" and "Total views" of the NBER working paper. Tables 17 and 18 present summary statistics and correlation coefficients for this sample.

TABLES 17 AND 18

The results of three negative binomial regressions are presented in table 19. None of the discussant characteristics are statistically significantly affecting any of the three outcome variables. The seeming discrepancy with the finding in 4.2 (where we do find a positive effect of discussant prolificness and citation count) is resolved by taking into account two limitations of RePEc citations: RePEc has a different coverage compared to Scopus (Scopus covers more sources

such as journals or books, while RePEc covers working papers) and we exclude citations to the working paper after it has been published in a journal.

TABLE 19 ABOUT HERE

Overall the results in this section lend some support for our general finding. Differences are likely due to the different nature of the dependent variables and the change a paper undergoes between pre-publication and journal publication.

C Tables

Table 1: Presentations in NBER Summer Institutes of financial NBER Working groups, by group and year.

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	total	share
AMRE		4 (5)	5 (6)									9 (11)	82%
AP	8 (8)	5 (8)	8 (9)	7 (8)	5 (8)	8 (9)	8 (9)	7 (9)	7 (7)	7 (8)	8 (9)	78 (92)	85%
CF	9 (10)	4 (7)	6 (9)	8 (10)	10 (13)	12 (14)	10 (11)	10 (13)	16 (18)	5 (6)	13 (14)	103 (125)	82%
EFCE	11 (15)	5 (16)	8 (16)	7 (14)	12 (17)	7 (15)	11 (15)	11 (15)	7 (15)	7 (15)	7 (15)	93 (168)	55%
EFDW	11 (14)											11 (14)	79%
EFEL		12 (16)	12 (15)	8 (12)	5 (11)	9 (11)	9 (12)	8 (12)	8 (10)	8 (12)	5 (10)	84 (121)	69%
EFWW		8 (12)	8 (12)	8 (12)	11 (12)	9 (12)	6 (12)	9 (12)	5 (12)	8 (12)	8 (12)	80 (120)	67%
HF											4 (6)	4 (6)	67%
IFM	6 (8)	6 (8)	10 (12)	8 (11)	9 (12)	5 (8)	7 (10)	7 (10)	8 (12)	11 (14)	9 (15)	86 (120)	72%
ME	8 (10)	5 (8)	6 (11)	8 (9)	8 (11)	5 (12)	12 (14)	5 (12)	7 (13)	9 (11)	9 (13)	82 (124)	66%
PERE	4 (5)	6 (6)	4 (6)	1 (6)	6 (9)	7 (10)	3 (10)	6 (10)	3 (11)	6 (9)	5 (6)	51 (88)	57%
RISK							3 (3)	1 (2)	2 (3)	9 (12)	6 (7)	21 (27)	78%
total	57 (70)	55 (86)	67 (96)	55 (82)	66 (93)	62 (91)	69 (96)	64 (95)	63 (101)	70 (99)	74 (107)	702 (1016)	
share	81%	64%	70%	67%	71%	68%	72%	67%	62%	71%	69%		69%

Notes: Table lists the presentations and the number of published papers in the NBER sample by year and NBER group. The NBER groups are Asset Marketing/Real Estate (AMRE), Asset Pricing (AP), Corporate Finance (CF), Impulse and Propagation Mechanisms (EFCE), Capital Markets in the Economy (EFEL), Forecasting & Empirical Methods in Macro & Finance (EFWW), Household Finance (HF), International Finance & Macroeconomics (IFM), Monetary Economics (ME), Finance & Macro (MEFM), Economics of Real Estate & Local Public Finance (PERE) and Risk of Financial Institutions (RISK). The first is the number of manuscripts that resulted in a publication until January 2019 and the number in brackets is the number of discussed manuscripts. For example, during the 2001 AP Summer Institute, 8 manuscripts were discussed of which 5 were eventually published in a Scopus-indexed journal. Column and row "share" indicate the share of publications to total presentations.

Table 2: Results of logistic and negative binomial estimations for assortative matching of discrete and real discussant characteristics to author and paper characteristics.

	logistic		*negative binomial*	
	Has discussion	Dis. experience	Dis. degree (coauthor)	Dis. degree (informal)
	(1)	(2)	(3)	(4)
Auth. total Euclid	−0.0001	0.00003	−0.0001	0.00001
	$p = 0.768$	$p = 0.698$	$p = 0.532$	$p = 0.935$
Auth. total experience	0.0004	0.008	0.016**	−0.001
	$p = 0.152$	$p = 0.211$	$p = 0.021$	$p = 0.914$
Auth. total experience^2		−0.00003	−0.0003**	−0.0001
		$p = 0.753$	$p = 0.024$	$p = 0.685$
# of authors	0.536***	−0.125**	−0.008	0.185***
	$p = 0.00003$	$p = 0.027$	$p = 0.888$	$p = 0.007$
Constant	−0.549			
	$p = 0.145$			
Discussion year-fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
NBER group-fixed effects	*No*	*Yes*	*Yes*	*Yes*
N	615	364	291	299
R^2	0.081			
Log Likelihood		-1,229.715	-689.591	-1,265.022
Akaike Inf. Crit.		2,503.430	1,423.181	2,574.045

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. *Has discussion* equals 1 if the paper has been discussed at an NBER Summer Institute. *Dis. experience* is the discussant's number of years between the year of the discussion and their first publication (negative values set to missing). *Dis. degree (coauthor)* is the discussant's degree in the giant component of the co-author network G to the year of the discussion. *Dis. degree (informal)* is the discussant's degree in the giant component of the network of informal collaboration H to the year of the discussion. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. *# of authors* is the number of authors of the published paper. Only papers considered that were presented exactly once at a Finance-related NBER Summer Institute and published until January 2019.

Table 3: Results of OLS estimations for assortative matching of continuous discussant characteristics to author and paper characteristics.

	Dis. Euclid	Dis. neighborhood (coauthor)	Dis. neighborhood (informal)
	(1)	(2)	(3)
Auth. total Euclid	−0.017	0.160	0.008
	$p = 0.557$	$p = 0.378$	$p = 0.708$
Auth. total experience	−0.767	−3.903	−1.916
	$p = 0.729$	$p = 0.653$	$p = 0.259$
Auth. total experience^2	0.065*	0.093	0.013
	$p = 0.081$	$p = 0.597$	$p = 0.634$
# of authors	−11.238	106.701	7.182
	$p = 0.564$	$p = 0.198$	$p = 0.625$
Constant	−23.643	2,868.420***	2,756.876***
	$p = 0.840$	$p = 0.00000$	$p = 0.000$
Discussion year-fixed effects	*Yes*	*Yes*	*Yes*
NBER group-fixed effects	*Yes*	*Yes*	*Yes*
N	364	158	298
R^2	0.196	0.919	0.979
Adjusted R^2	0.147	0.909	0.977

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. *Dis. Euclid* is the discussant's Euclidean index of citations (equation (1)) in the year of the discussion. *Dis. neighborhood (coauthor)* is the discussant's neighborhood centrality (equation (2)) in the giant component of the co-author network G corresponding to the year of the discussion. Same logic applies to *Dis. neighborhood (informal)*, which is computed in the giant component of network of informal collaboration H in the year of the discussion. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. *# of authors* is the number of authors of the published paper. Only papers considered that were presented exactly once at a Finance-related NBER Summer Institute and published until January 2019.

Table 4: Summary statistics for all variables in the discussants sample.

	N	Mean	Median	Std.Dev.	Min	Max
Academic success						
Total citations	615	105.4	55	154.92	0	1260
Top publication	615	0.5	1	0.50	0	1
Journal Impact Factor	596	8.3	7	5.35	0	23
Paper characteristics						
# of pages	615	29.5	29	11.89	1	87
# of authors	615	2.2	2	0.80	1	5
Age	615	3.4	3	1.85	0	13
Age^2	615	14.9	9	16.87	0	169
Author characteristics						
Auth. total Euclid	615	370.6	174	648.38	0	7965
Auth. total experience	615	21.0	18	16.90	−1	89
Auth. total $experience^2$	615	726.5	324	1054.55	0	7921
Treatment variable						
Has discussion	615	0.6	1	0.49	0	1

Notes: Total Citations is the total citation count until January 2019 according to Scopus. *Top publication* is a binary variable indicating publication in a Top 5 Econ or Top 3 Finance journal. *Journal Impact Factor* is the Journal Impact Factor in the year of publication according to Scimago. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. Age^2 is its square. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total* $experience^2$ is its square. *Has discussion* equals 1 if the paper had a discussant once during a Summer Institute of a Finance-related NBER working group between 2001 and 2010, and 0 otherwise.

Table 5: Spearman and Pearson correlation coefficients for all variables in the discussants sample.

Academic success											
Total citations		1.00	0.33	0.26	−0.38	0.04	0.15	0.28	−0.04	1.00	0.15
Top publication	0.25		0.21	0.09	0.24	0.33	−0.13	0.21	0.17	0.09	0.17
Journal Impact Factor	0.12	0.75		−0.38	0.12	1.00	−0.13	0.81	0.04	−0.08	0.09
Paper characteristics											
# of pages	0.21	0.40	0.45		0.13	0.81	0.10	1.00	0.04	−0.08	1.00
# of authors	0.06	0.14	0.16	0.09		0.43	−0.01	0.43	0.24	0.12	0.03
Age	−0.23	−0.15	−0.08	−0.08	0.03		−0.01	0.17	0.26	−0.01	0.03
Age^2	−0.21	−0.17	−0.11	−0.08	0.04	0.94		−0.04	0.43	−0.02	0.41
Author characteristics											
Auth. total Euclid	0.11	0.07	0.16	0.03	0.26	−0.05	−0.05		0.43	0.16	0.46
Auth. total experience	0.06	0.00	0.07	−0.02	0.48	−0.01	−0.03	0.58		0.09	0.45
Auth. total experience^2	0.03	0.01	0.08	−0.02	0.43	0.03	0.02	0.55	0.93		0.13
Treatment variable											
Has discussion	0.01	0.28	0.24	0.15	0.12	0.10	0.10	−0.03	−0.02	−0.01	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total Citations* is the total citation count until January 2019 according to Scopus. *Top publication* is a binary variable indicating publication in a Top 5 Econ or Top 3 Finance journal. *Journal Impact Factor* is the Journal Impact Factor in the year of publication according to Scimago. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. *Age*^{2} is its square. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. *Has discussion* equals 1 if the paper was discussed at least once during a Summer Institute of a Finance-related NBER working group between 2001 and 2010.

Table 6: Estimation results for the discussants sample.

	Total citations	Top publication	Journal Impact Factor
	negative binomial	*logistic*	*OLS*
	(1)	(2)	(3)
Auth. total Euclid	0.0004^{***}	0.0003^{*}	0.001^{***}
	$p = 0.00000$	$p = 0.070$	$p = 0.0001$
Auth. total experience	-0.005	-0.031^{*}	-0.069^{**}
	$p = 0.497$	$p = 0.052$	$p = 0.013$
Auth. total experience^2	-0.00003	0.0003	0.001^{*}
	$p = 0.773$	$p = 0.249$	$p = 0.053$
Has discussion	0.096	1.126^{***}	1.923^{***}
	$p = 0.213$	$p = 0.000$	$p = 0.004$
Constant	3.270^{***}	-3.843^{***}	-0.678
	$p = 0.000$	$p = 0.000$	$p = 0.495$
Paper characteristics	*Yes*	*Yes*	*Yes*
Publication year-FE	*Yes*	*No*	*No*
N	615	615	596
R^2			0.290
Adjusted R^2			0.281
Akaike Inf. Crit.	6,625.580	683.917	

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Standard errors clustered around NBER working group. Reported coefficients in columns (1) and (2) are marginal effects. *Total Citations* is the total citation count until January 2019 according to Scopus. *Top publication* is a binary variable indicating publication in a Top 5 Econ or Top 3 Finance journal. *Journal Impact Factor* is the Journal Impact Factor in the year of publication according to Scimago. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. *Age*^{2} is its square. *Auth. total Euclid* is the author's Euclidean index of citations (equation (1)) in the year before publication, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. *Has discussion* equals 1 if the paper was discussed at least once during a Summer Institute of a Finance-related NBER working group between 2001 and 2010.

Table 7: Summary statistics for all variables in the discussants characteristics sample.

	N	Mean	Median	Std.Dev.	Min	Max
Academic success						
Total citations	378	106.4	56	154.60	0	1260
Top publication	378	0.6	1	0.48	0	1
Journal Impact Factor	369	9.3	9	5.55	0	23
Paper characteristics						
# of pages	378	30.9	31	11.44	1	84
# of authors	378	2.3	2	0.79	1	5
Age	378	3.5	3	1.93	0	13
Age^2	378	16.2	9	18.71	0	169
Author characteristics						
Auth. total Euclid	378	354.6	174	634.61	0	7965
Auth. total experience	378	20.7	18	16.99	−1	89
Auth. total $experience^2$	378	716.4	306	1068.73	0	7921
Discussant characteristics						
Dis. Euclid	378	186.5	97	260.82	0	1670
Dis. experience	378	11.4	10	8.41	−5	38
Dis. $experience^2$	378	199.5	100	252.98	0	1444

Notes: Total Citations is the total citation count until January 2019 according to Scopus. *Top publication* is a binary variable indicating publication in a Top 5 Econ or Top 3 Finance journal. *Journal Impact Factor* is the Journal Impact Factor in the year of publication according to Scimago. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. Age^2 is its square. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. $Auth.\ total\ experience^2$ is its square. Same logic applies to discussant (*Dis.*) variables, which were counted in the year of the discussion.

Table 8: Spearman and Pearson correlation coefficients for all variables in the discussants characteristics sample.

Academic success												
Total citations		1.00	0.37	0.21	−0.40	0.02	0.80	0.01	1.00	0.12	−0.08	0.14	−0.23
Top publication	0.25		0.23	0.04	0.19	−0.01	0.44	0.01	0.48	0.04	0.18	1.00	−0.17
Journal Impact Factor	0.12	0.75		−0.40	0.10	−0.01	0.13	0.12	0.17	0.04	0.03	0.01	−0.08
Paper characteristics													
# of pages	0.15	0.42	0.49		0.10	0.37	−0.23	0.03	−0.17	0.21	0.03	0.01	0.01
# of authors	0.00	0.13	0.16	0.15		1.00	−0.23	0.03	−0.17	0.44	0.07	0.45	1.00
Age	−0.26	−0.26	−0.19	−0.10	0.00		0.14	0.23	0.21	0.48	−0.01	0.49	1.00
Age^2	−0.23	−0.27	−0.21	−0.11	0.00	0.94		0.80	0.07	1.00	−0.02	0.48	−0.11
Author characteristics													
Auth. total Euclid	0.06	0.14	0.21	0.06	0.23	−0.08	−0.06		0.07	0.14	0.04	0.02	−0.03
Auth. total experience	0.02	0.03	0.09	0.02	0.48	−0.01	−0.02	0.55		−0.08	0.13	−0.04	−0.04
Auth. total $experience^2$	−0.02	0.04	0.09	0.00	0.41	0.03	0.02	0.50	0.92		0.17	−0.04	−0.07
Discussant characteristics													
Dis. Euclid	−0.01	0.17	0.17	0.06	0.08	−0.05	−0.04	0.14	0.12	0.16		−0.40	−0.03
Dis. experience	−0.04	0.01	0.02	−0.04	−0.05	−0.01	0.00	0.17	0.08	0.08	0.57		−0.04
Dis. $experience^2$	−0.06	−0.01	0.01	−0.05	−0.06	0.02	0.02	0.16	0.07	0.08	0.50	0.94	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total Citations* is the total citation count until January 2019 according to Scopus. *Top publication* is a binary variable indicating publication in a Top 5 Econ or Top 3 Finance journal. *Journal Impact Factor* is the Journal Impact Factor in the year of publication according to Scimago. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. Age^2 is its square. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total* $experience^2$ is its square. Same logic applies to discussant (*Dis.*) variables which were counted in the year of the discussion.

Table 9: Estimation results for the discussants characteristics sample.

	Total citations	Top publication	Journal Impact Factor
	negative binomial	*logistic*	*OLS*
	(1)	(2)	(3)
Auth. total Euclid	0.0002**	0.001	0.001***
	$p = 0.016$	$p = 0.102$	$p = 0.0005$
Auth. total experience	−0.001	−0.039	−0.041
	$p = 0.881$	$p = 0.110$	$p = 0.304$
Auth. total experience^2	−0.00004	0.0004	0.0004
	$p = 0.737$	$p = 0.308$	$p = 0.500$
Dis. Euclid	0.001***	0.001	0.002**
	$p = 0.0005$	$p = 0.126$	$p = 0.024$
Dis. experience	0.020	−0.030	−0.040
	$p = 0.231$	$p = 0.524$	$p = 0.625$
Dis. experience^2	−0.001**	0.001	0.001
	$p = 0.046$	$p = 0.634$	$p = 0.695$
Constant	3.397***	−3.868***	
	$p = 0.000$	$p = 0.008$	
Paper characteristics	*Yes*	*Yes*	*Yes*
NBER group-FE	*Yes*	*Yes*	*Yes*
Publication year-FE	*Yes*	*No*	*No*
N	378	378	369
R^2			0.356
Adjusted R^2			0.324
Akaike Inf. Crit.	4,084.442	380.401	

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Standard errors clustered around NBER working group. Reported coefficients in columns (1) and (2) are marginal effects. *Total Citations* is the total citation count until January 2019 according to Scopus. *Top publication* is a binary variable indicating publication in a Top 5 Econ or Top 3 Finance journal. *Journal Impact Factor* is the Journal Impact Factor in the year of publication according to Scimago. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. Same logic applies to discussant (*Dis.*) variables which were counted in the year of the discussion. Paper characteristics include the number of pages, the number of authors, the number of years of between the paper's first presentation at an NBER Summer Institute and publication, as well as the square thereof.

Table 10: Summary statistics for all variables in the network position sample.

	N	Mean	Median	Std.Dev.	Min	Max
Academic success						
Total citations	345	107.7	57	158.04	0	1260
Paper characteristics						
# of pages	345	31.4	32	11.20	1	84
# of authors	345	2.3	2	0.79	1	5
Age	345	3.6	3	1.95	0	13
Age^2	345	16.4	9	19.06	0	169
Author characteristics						
Auth. total Euclid	345	369.8	181	657.06	0	7965
Auth. total experience	345	20.9	17	17.11	−1	89
Auth. total experience^2	345	728.6	289	1092.52	0	7921
Discussant characteristics						
Dis. Euclid	345	198.6	99	269.04	0	1670
Dis. experience	345	11.6	10	8.36	−5	38
Dis. experience^2	345	204.2	100	256.67	0	1444
Discussant co-author network						
Dis. degree (coauthor)	291	4.7	4	3.72	1	26
Dis. neighborhood (coauthor)	158	6289.2	6537	1994.10	2502	10247
Discussant informal network						
Dis. degree (informal)	299	31.7	23	28.72	1	161
Dis. neighborhood (informal)	298	4348.0	4344	1106.39	2271	6159

Notes: Total Citations is the total citation count until January 2019 according to Scopus. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. *Age*^{2} is its square. *Auth. total Euclid* is the authors' combined Euclidean index of citations in the year before publication (equation (1)). *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. Same logic applies to *Dis. Euclid*, *Dis. experience* and *Dis. experience*^{2}, which were counted in the year of the discussion. *Dis. degree (coauthor)* is the discussants' degree, and *Dis. neighborhood (coauthor)* the discussants' neighborhood centrality (equation (2)) in the giant component of the co-author network of informal collaboration G corresponding to the year of the discussion. Same logic applies to *Dis. degree (informal)* and *Dis. neighborhood (informal)*, which are computed in the giant component of network of informal collaboration H in the year of the discussion.

Table 11: Spearman and Pearson correlation coefficients for all variables in the network position sample.

Academic success															
Total citations		1.00	0.19	−0.41	0.10	−0.01	0.19	0.03	0.15	0.49	−0.41	−0.04	−0.02	−0.10	0.03
Paper characteristics															
# of pages	0.13		0.03	−0.41	0.10	0.06	1.00	0.02	0.09	0.48	−0.08	−0.07	−0.03	−0.10	0.49
# of authors	−0.01	0.11		0.19	0.01	−0.13	0.12	0.05	0.03	0.04	0.02	0.00	−0.09	1.00	−0.02
Age	−0.25	−0.10	0.01		−0.02	0.04	−0.08	−0.02	0.12	−0.02	1.00	−0.04	−0.03	0.87	−0.02
Age^2	−0.23	−0.11	0.01	0.94		−0.39	−0.08	−0.02	1.00	−0.02	1.00	−0.05	−0.04	0.87	0.87
Author characteristics															
Auth. total Euclid	0.05	0.03	0.21	−0.08	−0.06		0.17	0.15	0.02	0.11	−0.10	−0.41	−0.07	0.16	1.00
Auth. total experience	0.01	0.01	0.48	0.00	−0.02	0.56		0.17	0.02	0.15	−0.02	−0.08	0.00	0.10	1.00
Auth. total experience^2	−0.03	−0.01	0.41	0.03	0.02	0.51	0.92		0.44	0.21	−0.03	0.02	−0.04	0.10	0.10
Discussant characteristics															
Dis. Euclid	−0.01	0.05	0.08	−0.06	−0.05	0.13	0.12	0.16		0.16	−0.09	1.00	−0.05	0.07	0.09
Dis. experience	−0.04	−0.05	−0.04	0.00	0.00	0.18	0.08	0.08	0.57		−0.03	1.00	0.19	0.08	0.09
Dis. experience^2	−0.06	−0.07	−0.05	0.02	0.03	0.17	0.08	0.08	0.50	0.94		−0.10	0.17	0.11	0.03
Discussant co-author network															
Dis. degree (coauthor)	0.11	0.17	0.07	−0.04	−0.06	0.05	0.06	0.04	0.23	0.15	0.09		0.44	0.07	−0.02
Dis. neighborhood (coauthor)	−0.12	0.18	0.15	0.05	0.08	0.19	0.03	0.05	0.27	0.11	0.09	0.09		0.10	0.02
Discussant informal network															
Dis. degree (informal)	−0.03	0.09	0.18	−0.05	−0.08	0.04	0.00	0.01	0.36	0.09	0.04	0.28	0.26		−0.01
Dis. neighborhood (informal)	−0.32	0.12	0.15	−0.07	−0.07	0.16	0.03	0.07	0.19	0.14	0.12	0.07	0.91	0.26	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total Citations* is the total citation count until January 2019 according to Scopus. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *# of pages* and *# of authors* is the simple count of pages and authors, respectively. *Age* is the number of years between the paper's first presentation at an NBER Summer Institute and the publication year. *Age*^{2} is its square. *Auth. total Euclid* is the author's Euclidean index of citations (equation (1)) in the year before publication, summed over all authors. *Auth. total experience* is the combined number of years between the year before publication and the publication year of each author's first paper. *Auth. total experience*^{2} is its square. Same logic applies to *Dis. Euclid*, *Dis. experience* and *Dis. experience*^{2}. *Dis. degree* is the discussants' degree, and *Dis. neighborhood* the discussants' neighborhood centrality (equation (2)) in the giant component of the co-author network G corresponding to the year of the discussion. Same logic applies to *Dis. degree (informal)* and *Dis. neighborhood (informal)*, which are computed in the giant component of network of informal collaboration H in the year of the discussion.

Table 12: Results of negative binomial regression for Total Citations on discussant's coauthor network position.

	Total citations	
	(1)	(2)
Dis. Euclid	0.001^{***}	-0.001
	$p = 0.0003$	$p = 0.166$
Dis. experience	0.005	0.070^{**}
	$p = 0.814$	$p = 0.019$
Dis. experience^2	-0.001	-0.002^{**}
	$p = 0.320$	$p = 0.025$
Dis. degree (coauthor)	0.022	
	$p = 0.136$	
Dis. neighborhood (coauthor)		0.0001
		$p = 0.314$
Constant	5.020^{***}	3.720^{***}
	$p = 0.000$	$p = 0.003$
Paper characteristics	*Yes*	*Yes*
Author characteristics	*Yes*	*Yes*
Publication year-FE	*Yes*	*Yes*
NBER group-FE	*Yes*	*Yes*
N	291	158
Akaike Inf. Crit.	3,117.000	1,804.000

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Standard errors clustered around NBER working group. Reported coefficients are marginal effects. *Total Citations* is the total citation count until January 2019 according to Scopus. *# of discussants* is the count of distinct discussants that discussed the manuscript. *Dis. Euclid* is the discussants' Euclidean index of citations (equation (1)) in the year of the discussion. *Dis. experience* is the number of years between the year of discussion and the publication year of the first paper. *Dis. experience*^{2} is its square. *Dis. degree (coauthor)* is the discussants' degree and *Dis. neighborhood (coauthor)* the discussants' neighborhood centrality (equation (2)) in the giant component of the coauthor network G corresponding to the year of the discussion. Paper characteristics include the number of pages, the number of authors, the number of years of between the paper's first presentation at an NBER Summer Institute and publication, as well as the square thereof. Author controls include the authors' combined Euclidean index of citations in the year before publication, their combined experience, and the square thereof.

Table 13: Results of negative binomial regression for Total Citations on discussant's informal network position.

	Total citations	
	(1)	(2)
Dis. Euclid	0.001^{***}	0.001^{***}
	$p = 0.00004$	$p = 0.00005$
Dis. experience	0.001	−0.002
	$p = 0.955$	$p = 0.940$
Dis. experience^2	−0.001	−0.0004
	$p = 0.425$	$p = 0.543$
Dis. degree (informal)	−0.002	
	$p = 0.464$	
Dis. neighborhood (informal)		0.0003
		$p = 0.331$
Constant	4.170^{***}	2.860^{**}
	$p = 0.000$	$p = 0.046$
Paper characteristics	*Yes*	*Yes*
Author characteristics	*Yes*	*Yes*
Publication year-FE	*Yes*	*Yes*
NBER group-FE	*Yes*	*Yes*
N	299	298
Akaike Inf. Crit.	3,254.000	3,245.000

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Standard errors clustered around NBER working group. Reported coefficients are marginal effects. *Total Citations* is the total citation count until January 2019 according to Scopus. *# of discussants* is the count of distinct discussants that discussed the manuscript. *Dis. Euclid* is the discussants' Euclidean index of citations (equation (1)) in the year of the discussion. *Dis. experience* is the number of years between the year of discussion and the publication year of the first paper. *Dis. experience*^{2} is its square. *Dis. degree (informal)* is the discussants' degree and *Dis. neighborhood (informal)* the discussants' neighborhood centrality (equation (2)) in the giant component of the network of informal collaboration H corresponding to the year of the discussion. Paper characteristics include the number of pages, the number of authors, the number of years of between the paper's first presentation at an NBER Summer Institute and publication, as well as the square thereof. Author controls include the authors' combined Euclidean index of citations in the year before publication, their combined experience, and the square thereof.

Table 14: Summary statistics for all variables in the RePEc discussants sample.

	N	Mean	Median	Std.Dev.	Min	Max
Academic success						
Total citations (RePEc)	236	10.5	5	16.99	0	146
Total downloads	235	79.7	52	83.37	1	512
Total views	235	213.5	146	209.23	3	1126
Paper characteristics						
# of authors	236	2.3	2	0.73	1	4
Author characteristics						
Auth. total Euclid (WP)	236	431.6	237	708.49	0	7965
Auth. total experience (WP)	236	21.8	20	16.52	−7	73
Auth. total experience^2 (WP)	236	746.6	400	1013.94	0	5329
Treatment variable						
Has discussion	236	0.7	1	0.47	0	1

Notes: Total Citations (RePEc) is the total citation count until January 2019 according to RePEc. *Total downloads* is the total number of downloads until January 2019. *Total views* is the total number of views until January 2019. *# of authors* is the simple count of authors. *Auth. total Euclid (WP)* is the authors' combined Euclidean index of citations in the year of working paper publication (equation (1)). *Auth. total experience (WP)* is the combined number of years between the year of working paper publication and the publication year of each author's first paper. *Auth. total experience^2 (WP)* is its square. *Has discussion* equals 1 if the paper had a discussant once during a Summer Institute of a Finance-related NBER working group between 2001 and 2010, and 0 otherwise.

Table 15: Spearman and Pearson correlation coefficients for all variables in the RePEc discussants sample.

Academic success								
Total citations (RePEc)		1.00	0.58	0.02	0.05	0.86	−0.06	0.03
Total downloads	0.42		0.59	0.10	−0.24	0.05	0.59	0.03
Total views	0.43	0.90		0.05	0.58	0.22	0.86	−0.02
Paper characteristics								
# of authors	0.05	0.09	0.06		1.00	0.14	1.00	0.02
Author characteristics								
Auth. total Euclid (WP)	0.05	0.14	0.05	0.23		0.14	0.02	0.05
Auth. total experience (WP)	0.13	0.10	0.03	0.42	0.53		0.08	0.02
Auth. total experience^2 (WP)	0.13	0.07	0.01	0.41	0.49	0.94		1.00
Treatment variable								
Has discussion	−0.15	−0.09	−0.01	0.08	−0.06	−0.08	−0.08	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total Citations (RePEc)* is the total citation count until January 2019 according to RePEc. *Total downloads* is the total number of downloads until January 2019. *Total views* is the total number of views until January 2019. *# of authors* is the simple count of authors. *Auth. total Euclid (WP)* is the authors' combined Euclidean index of citations in the year of working paper publication (equation (1)). *Auth. total experience (WP)* is the combined number of years between the year of working paper publication and the publication year of each author's first paper. *Auth. total experience^2 (WP)* is its square. *Has discussion* equals 1 if the paper was discussed at least once during a Summer Institute of a Finance-related NBER working group between 2001 and 2010.

Table 16: Estimation results for the RePEc discussants sample.

	Total citations (RePEc)	Total downloads	Total views
	(1)	(2)	(3)
Auth. total Euclid	−0.00002	0.0002***	0.0001
	$p = 0.914$	$p = 0.007$	$p = 0.132$
Auth. total experience	0.025	0.016*	0.010
	$p = 0.101$	$p = 0.099$	$p = 0.310$
Auth. total experience^2	−0.0003	−0.0003*	−0.0002
	$p = 0.246$	$p = 0.086$	$p = 0.193$
Has discussion	−0.562***	−0.193*	−0.111
	$p = 0.002$	$p = 0.090$	$p = 0.343$
Constant	2.491***	4.095***	4.999***
	$p = 0.000$	$p = 0.000$	$p = 0.000$
Paper characteristics	*Yes*	*Yes*	*Yes*
WP publication year-FE	*Yes*	*No*	*No*
N	236	235	235
Akaike Inf. Crit.	1,579.869	2,513.682	2,984.863

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Standard errors clustered around NBER working group. Reported coefficients are marginal effects. *Total Citations (RePEc)* is the total citation count until January 2019 according to RePEc. *Total downloads* is the total number of downloads until January 2019. *Total views* is the total number of views until January 2019. *# of authors* is the simple count of authors. *Auth. total Euclid (WP)* is the authors' combined Euclidean index of citations in the year of working paper publication (equation (1)). *Auth. total experience (WP)* is the combined number of years between the year of working paper publication and the publication year of each author's first paper. *Auth. total experience^2 (WP)* is its square. *Has discussion* equals 1 if the paper was discussed at least once during a Summer Institute of a Finance-related NBER working group between 2001 and 2010.

Table 17: Summary statistics for all variables in the RePEc discussants characteristics sample.

	N	Mean	Median	Std.Dev.	Min	Max
Academic success						
Total citations (RePEc)	157	8.7	4	14.64	0	97
Total downloads	157	74.6	52	71.37	1	481
Total views	157	212.2	152	204.33	3	1126
Paper characteristics						
# of authors	157	2.4	2	0.74	1	4
Author characteristics						
Auth. total Euclid (WP)	157	403.5	202	760.43	0	7965
Auth. total experience (WP)	157	20.8	18	16.08	−7	73
Auth. total experience^2 (WP)	157	689.3	324	984.97	0	5329
Discussant characteristics						
Dis. Euclid	157	215.0	108	302.18	0	1670
Dis. experience	157	11.6	10	8.49	−5	35
Dis. experience^2	157	206.3	100	254.86	0	1225

Notes: Total Citations (RePEc) is the total citation count until January 2019 according to RePEc. *Total downloads* is the total number of downloads until January 2019. *Total views* is the total number of views until January 2019. *# of authors* is the simple count of authors. *Auth. total Euclid (WP)* is the authors' combined Euclidean index of citations in the year of working paper publication (equation (1)). *Auth. total experience (WP)* is the combined number of years between the year of working paper publication and the publication year of each author's first paper. *Auth. total experience^2 (WP)* is its square. Same logic applies to discussant (*Dis.*) variables, which were counted in the year of the discussion.

Table 18: Spearman and Pearson correlation coefficients for all variables in the RePEc discussants characteristics sample.

Academic success										
Total citations (RePEc)		1.00	0.60	−0.05	0.03	0.60	0.11	0.89	0.07	0.38
Total downloads	0.38		0.61	0.09	−0.15	1.00	0.11	1.00	0.06	0.00
Total views	0.44	0.88		0.03	−0.12	0.89	0.02	−0.06	−0.05	−0.06
Paper characteristics										
# of authors	−0.08	0.01	0.00		−0.12	−0.03	0.06	0.07	−0.03	−0.06
Author characteristics										
Auth. total Euclid (WP)	0.01	0.07	0.00	0.17		0.17	0.06	0.01	−0.06	0.09
Auth. total experience (WP)	0.08	0.06	0.01	0.41	0.49		0.61	0.02	1.00	0.17
Auth. total experience^2 (WP)	0.03	0.01	−0.03	0.38	0.45	0.94		0.02	0.35	0.07
Discussant characteristics										
Dis. Euclid	−0.13	0.05	0.01	0.03	0.08	0.08	0.11		0.38	0.35
Dis. experience	−0.14	0.08	0.10	−0.06	0.19	0.04	0.04	0.60		1.00
Dis. experience^2	−0.10	0.09	0.11	−0.06	0.19	0.06	0.05	0.56	0.94	

Notes: Upper triangular shows Spearman (rank) correlation coefficients, while lower triangular shows Pearson correlation coefficients. *Total Citations (RePEc)* is the total citation count until January 2019 according to RePEc. *Total downloads* is the total number of downloads until January 2019. *Total views* is the total number of views until January 2019. *# of authors* is the simple count of authors. *Auth. total Euclid (WP)* is the authors' combined Euclidean index of citations in the year of working paper publication (equation (1)). *Auth. total experience (WP)* is the combined number of years between the year of working paper publication and the publication year of each author's first paper. *Auth. total experience^2 (WP)* is its square. Same logic applies to discussant (*Dis.*) variables which were counted in the year of the discussion.

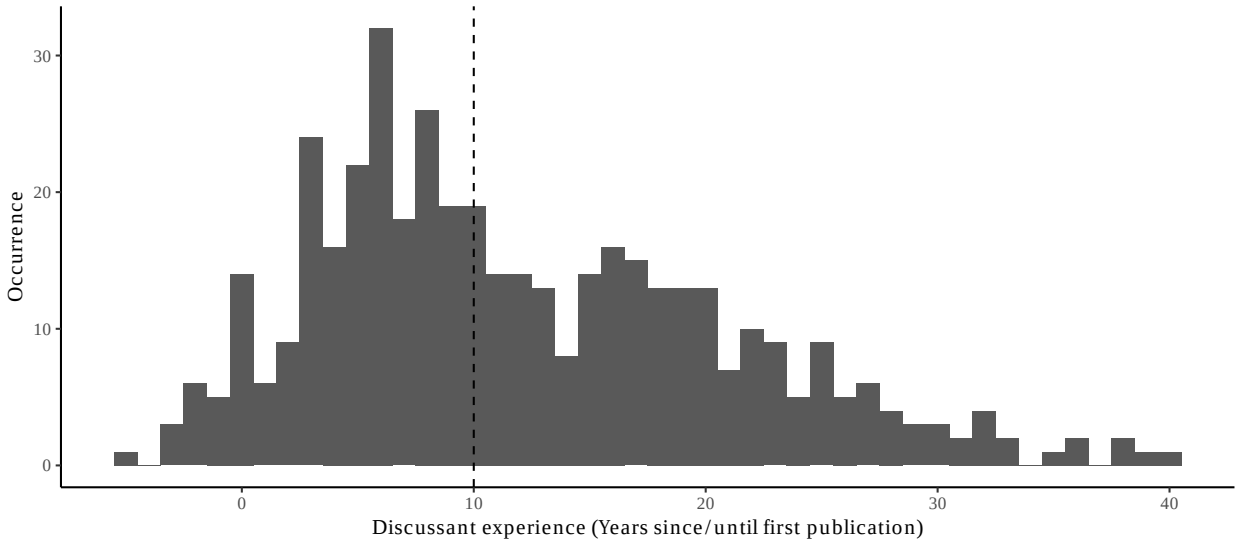
Table 19: Estimation results for the RePEc discussants characteristics sample.

	Total citations (RePEc)	Total downloads	Total views
	(1)	(2)	(3)
Auth. total Euclid (WP)	−0.040	−0.052	−0.026
	$p = 0.814$	$p = 0.600$	$p = 0.803$
Auth. total experience (WP)	0.0005***	0.0002	0.0001
	$p = 0.006$	$p = 0.102$	$p = 0.285$
Auth. total experience^2 (WP)	−0.014	0.017	0.007
	$p = 0.507$	$p = 0.187$	$p = 0.615$
Dis. Euclid	0.00001	−0.0002	−0.0001
	$p = 0.974$	$p = 0.256$	$p = 0.707$
Dis. experience	0.0001	0.0003	0.0002
	$p = 0.838$	$p = 0.314$	$p = 0.596$
Dis. experience^2	−0.058	−0.004	0.003
	$p = 0.134$	$p = 0.862$	$p = 0.904$
experiencesq_dis	0.001	−0.00002	−0.0002
	$p = 0.440$	$p = 0.980$	$p = 0.831$
Constant	4.497***	5.114***	5.648***
	$p = 0.004$	$p = 0.00000$	$p = 0.000$
Paper characteristics	*Yes*	*Yes*	*Yes*
NBER Group-FE	*Yes*	*Yes*	*Yes*
WP publication year-FE	*Yes*	*Yes*	*Yes*
N	157	157	157
Akaike Inf. Crit.	995.468	1,670.676	1,996.178

Notes: ***, ** and * indicate statistical significance to the 1, 5 and 10 percent level, respectively. Standard errors clustered around NBER working group. Reported coefficients are marginal effects. *Total Citations (RePEc)* is the total citation count until January 2019 according to RePEc. *Total downloads* is the total number of downloads until January 2019. *Total views* is the total number of views until January 2019. *# of authors* is the simple count of authors. *Auth. total Euclid (WP)* is the authors' combined Euclidean index of citations in the year of working paper publication (equation (1)). *Auth. total experience (WP)* is the combined number of years between the year of working paper publication and the publication year of each author's first paper. *Auth. total experience^2 (WP)* is its square. Same logic applies to discussant (*Dis.*) variables which were counted in the year of the discussion.

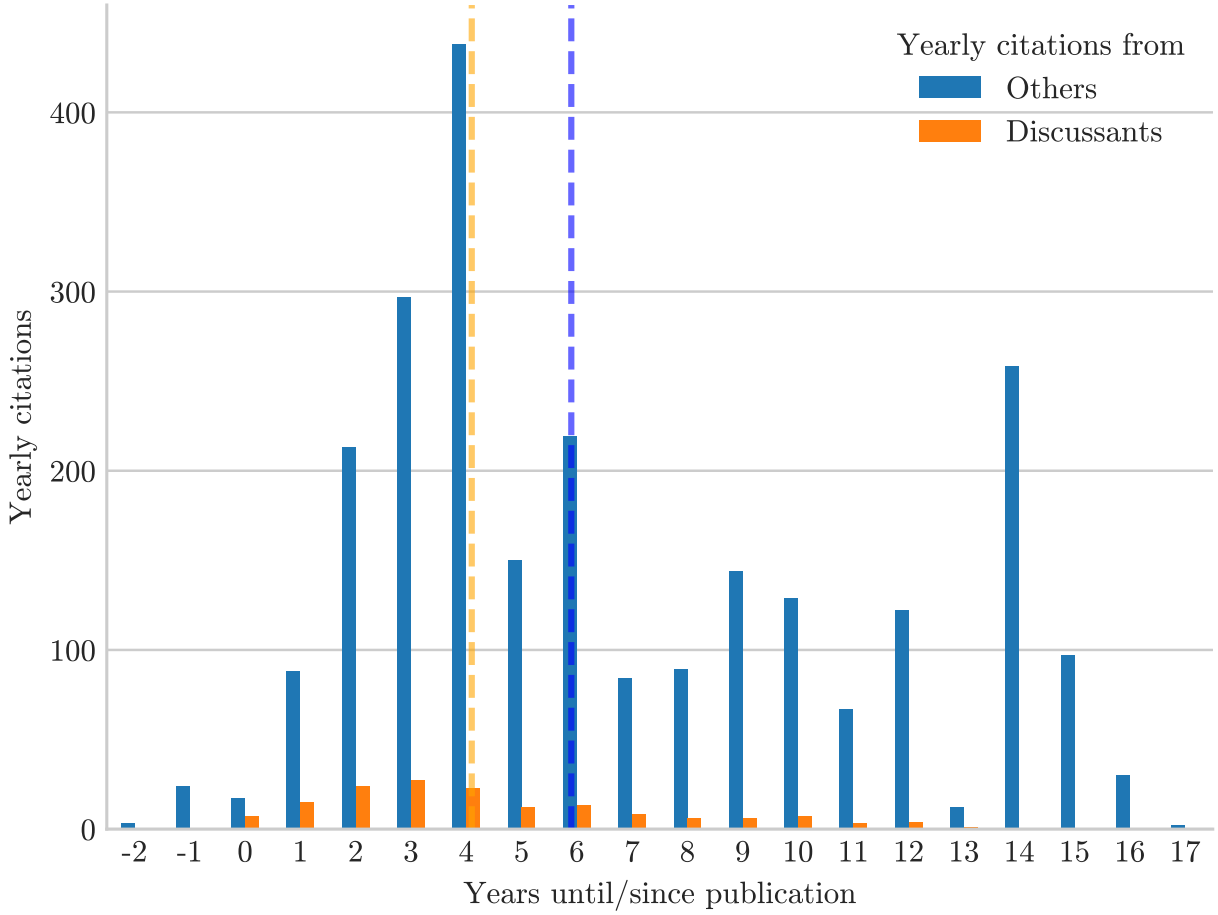
D Figures

Figure 1: Histogram for discussant's academic experience in year of discussion it.



Notes: Figure shows the histogram of discussant's experience in the year of the discussion. Experience is measures as the number of years since/until the first recorded publication in Scopus. The dashed line indicates the median value (10 years).

Figure 2: Barplot showing when papers are cited by its discussants and others, by number of how many commenters cite it.



Notes: Figure indicates the time lag before/since publication discussants cite the paper they discussed, as compared to non-discussants that cite the paper. Dashed lines indicate the lag until which 50% of the citations from that group occur. The total number of papers is 5320. Citations from "Discussant" exclude citations from papers that are co-authored with the original authors. Citations from "Others" include self-citations.

References

Abadie, A., Athey, S., Imbens, G. and Wooldridge, J. (2017), 'When Should You Adjust Standard Errors for Clustering?', *National Bureau of Economic Research, Working Paper*.

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.

Bandiera, O. and Rasul, I. (2006), 'Social Networks and Technology Adoption in Northern Mozambique', *The Economic Journal* **116**(514), 869–902.

Banerjee, A., Chandrasekhar, A. G., Duflo, E. and Jackson, M. O. (2013), 'The Diffusion of Microfinance', *Science* **341**(6144), 1236498–1236498.

Beaman, L., Benyishay, A., Magruder, J. and Mobarak, A. M. (2018), 'Can Network Theory-based Targeting Increase Technology Adoption?', *NBER Working Paper Series* **24912**.

Borjas, G. J. and Doran, K. B. (2015), 'Which Peers Matter? The Relative Impacts of Collaborators, Colleagues, and Competitors', *Review of Economics and Statistics* **97**(5), 11041117.

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

Cai, J., Janvry, A. D. and Sadoulet, E. (2015), 'Social Networks and the Decision to Insure', *American Economic Journal: Applied Economics* **7**(2), 81–108.

Combes, P.-p. and Linnemer, L. (2010), 'Inferring missing citations: a quantitative multi-criteria ranking of all jornals in economics', *HAL, Working Paper*.

Conley, T. G. and Udry, C. R. (2010), 'Learning about a New Technology: Pineapple in Ghana', *The American Economic Review* **100**(1), 35–69.

Coupé, T., Ginsburgh, V. and Noury, A. (2010), 'Are leading papers of better quality? Evidence from a natural experiment', *Oxford Economic Papers* **62**(1), 1–11.

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

Duflo, E. and Saez, E. (2003), 'The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment', *The Quarterly Journal of Economics* **118**(3), 815–842.

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

Feenberg, D., Ganguli, I., Gaulé, P. and Gruber, J. (2017), 'Its Good to Be First: Order Bias in Reading and Citing NBER Working Papers', *Review of Economics and Statistics* **99**(1), 32–39.

Gans, J. S. and Shepherd, G. B. (1994), 'How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists', *Journal of Economic Perspectives* **8**(1), 165–179.

Georg, C.-P. and Rose, M. E. (2018), 'What 5,000 Acknowledgements Tell Us About Informal Collaboration in Financial Economics', *Max Planck Institute for Innovation & Competition Discussion Paper 11*.

Golub, B. and Jackson, M. O. (2010), 'Naïve Learning in Social Networks and the Wisdom of Crowds', *American Economic Journal: Microeconomics* **2**(1), 112–149.

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

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

Heckman, J. and Moktan, S. (2018), 'Publishing and Promotion in Economics: The Tyranny of the Top Five', *NBER Working Paper*.

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

Laband, D. N. and Tollison, R. D. (2003), 'Good colleagues', *Journal of Economic Behavior & Organization* **52**(4), 505–512.

Merton, R. K. (1968), 'The Matthew Effect in Science: The reward and communication systems of science are considered', *Science* **159**(3810), 56–63.

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

Oswald, A. J. (2010), 'A suggested method for the measurement of world-leading research (illustrated with data on economics)', *Scientometrics* **84**(1), 99–113.

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

Ponomariov, B. and Boardman, C. (2016), 'What is co-authorship?', *Scientometrics* **109**(3), 1939–1963.

Richardson, M. and Domingos, P. (2002), Mining knowledge-sharing sites for viral marketing, *in* 'Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '02', ACM Press, New York, New York, USA, p. 61.

Rose, M. E. and Kitchin, J. R. (2019), 'scopus: Scriptable bibliometrics using a Python interface to Scopus', *Max Planck Institute for Innovation & Competition Research Paper* (19-03).

Waldinger, F. (2010), 'Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany', *Journal of Political Economy* **118**(4), 787–831.

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.