Collaboration and Power Laws for Financial Research Institutions[1]

Hui Dong[2]，Dan Luo[3]，Xudong Zeng[4] ，Zhentao Zou[5]

（School of Accountancy, Shanghai University of Finance and Economics, Shanghai, China; E-mail: dong.hui@mail.shufe.edu.cn；School of Finance, Shanghai University of Finance and Economics, Shanghai, China; and Shanghai Key Laboratory of Financial Information Technology, Shanghai, China; Email: luo.dan@mail.shufe.edu.cn.；School of Finance, Shanghai University of Finance and Economics, Shanghai, China; and Shanghai Key Laboratory of Financial Information Technology, Shanghai, China; E-mail: zeng.xudong@mail.shufe.edu.cn；Economics and Management School, Wuhan University, Wuhan, China; E-mail: zhentao zou@163.com.）

Abstract：A new means to promote collaboration is for one researcher to work across multiple institutions. We show that, accompanying the fast growth of cross-affiliation in financial research, scale-free power laws characterize the resulting highly-skewed distributions of top finance journal publications of worldwide institutions. We propose an explanation of the empirical power laws, based on a network model featuring two identified mechanisms: nonlinear growth and linear preferential attachment. The model indicates that preferential allocation of 87% of all publications effectively engenders dispersion in the research output of the institutions, while accelerated growth in collaboration provides a counterforce that restores their homogeneity.

Keywords: Research collaboration, Cross-affiliation, Power laws, Accelerated network, Preferential attachment

JEL Classifications: D85, G00

1. Introduction

Rapid technological progress empowers people with easier ways of communication and enhanced efficiency in performing research tasks. Such facilitation has greatly accelerated collaboration among researchers, as evidently witnessed in all branches of science including

[1] The authors are grateful for the valuable comments of Andrew Carverhill, Aaron Clauset, Xiaoji Lin, Weidong Tian, Liyan Yang, and seminar participants at Wuhan University. All data needed to evaluate the conclusions in the paper may be requested from the authors.

[2] School of Accountancy, Shanghai University of Finance and Economics, Shanghai, China; E-mail: dong.hui@mail.shufe.edu.cn.

[3] School of Finance, Shanghai University of Finance and Economics, Shanghai, China; and Shanghai Key Laboratory of Financial Information Technology, Shanghai, China; Email: luo.dan@mail.shufe.edu.cn.

[4] School of Finance, Shanghai University of Finance and Economics, Shanghai, China; and Shanghai Key Laboratory of Financial Information Technology, Shanghai, China; E-mail: zeng.xudong@mail.shufe.edu.cn.

[5] Economics and Management School, Wuhan University, Wuhan, China; E-mail: zhentao zou@163.com.

physics, mathematics, biology, neuroscience, computer science, and social sciences (see, e.g., Newman, 2001b,c,d, 2004; Barab´asi et al., 2002; Wuchty, Jones, and Uzzi, 2007). In the discipline of economics, in particular, the questions asked increasingly call upon concerted efforts of researchers with specific expertises. Card and DellaVigna (2013) survey top economics journals during 1970-2012 and demonstrate that the number of authors per paper has nearly doubled, and the researchers have submitted twice as many papers per year, rendering the top journals under the constant bombardments of submissions. The world is becoming a smaller village thanks to enlarging coauthorships (Goyal, van der Leij, and Moraga-Gonz´alez, 2006).

We observe a new form of cooperation in which researchers expand their scope of collaboration by affiliating simultaneously with multiple institutions. Top journal space is always limited.[6] The ability to assemble brainpower and other valuable resources (such as funding, data access, hardware, and visibility) across institutions is often vital for successfully publishing in top journals. Here, we focus on the three widely-acknowledged top journals in financial economics, namely, the *Journal of Finance* (JF, founded in 1946), the *Journal of Financial Economics* (JFE, founded in 1974), and the *Review of Financial Studies* (RFS, founded in 1986), which serve as ideal exemplifications.

To visualize the pattern of collaboration in financial research, we present in Panel A of Figure 1 the average number of coauthorships from 1980 to 2016 by aggregating the three journals. Coauthorships display steady, approximately linear growth observed earlier in other fields, such as economics (Card and DellaVigna, 2013) and neuroscience (Baraba´si et al., 2002). The new form of collaboration is shown in Panel B of Figure 1. This figure presents the average number of affiliated institutions reported per author by aggregating the three journals.[7] The number lingered at a low level before 1995 but took off thereafter, that is, in a little more than two decades, the average number of institutions per author increased steadily from below 1.1 to above 1.3. Put differently, on average, 1 out of 9 authors was affiliated with multiple institutions before 1995, and the ratio had escalated to 1 out of 3 by 2016. Panel C of Figure 1

[6] The total number of papers published by the top economics journals actually decreased slightly accompanying the outburst of submissions. As a result, the acceptance ratios have fallen off dramatically. See Card and DellaVigna (2013).

[7] Prior to 1980, an author commonly reported a single affiliation except in rare cases.

reports the average number of institutions per author separately for each journal. The same trend is prevalent across all three.

Top journal publications have a powerful influence on the impact of economic research. They also carry vast weight in the career paths of academic economists, the allocation of research funding, and the hiring decisions of research institutions. The new form of collaboration through cross-affiliation thus imparts fresh insights into the growth of research institutions gauged by their top journal publication records. In this study, we show that research institutions self-organize into a scale-free state and their top finance journal publications are characterized by power laws in the upper tail. The upper tail retains merely 6.0% of the totally 828 institutions at the end of 2016, while these most prolific institutions produce 61.3% of the totally 13548 top journal publications till the end of 2016. The power laws govern institutions of diverse nature, scattered across geographic region and time of establishment. They are also robust over the last six years and when considered separately for each journal.

Our results provide an important regularity for theories of research institution growth, which must target the empirical power laws documented here. We propose a model of network growth that explains the scale invariance, based on further examinations into the structural reasons for the emergence of the scale-free properties. Specifically, a research institution enters the research collaboration network as a node when it first publishes on a top finance journal and its degree henceforth equals its total number of publications. Two institutions are linked if they are affiliated to by one or more authors of a published article in a top journal. The deep collaboration in financial research is revealed by a giant component of the network, which consists of 99.5% of all the institutions at the end of 2016. The dynamics of the system are driven by nonlinear growth through enlarging coauthorships and cross-affiliation, as well as linear preferential attachment by incrementing total publications of institutions at a rate proportional to their current counts. The two mechanisms are generic, regardless of the types (universities, research centers, banks, etc.) and other characteristics (location, time of establishment, etc.) of the institutions.

Our study, to the best of our knowledge, is the first to explore collaboration on the institution level, albeit the extensive literature of collaboration networks on the individual researcher level (see, e.g., Newman, 2001b,c,d, 2004; Goyal, van der Leij, and Moraga-Gonz´alez, 2006). We characterize the structure of knowledge, collaboration patterns, and growth mechanism of the collaboration network for financial research institutions. Our empirical investigation reveals the scale-free and nonlinear growth features of the collaboration network, and our model enables a quantitative understanding of the forces shaping the allocation of top journal publications. We shows that cumulative advantages constitute a dominant channel for the allocation of top finance journal publications. In particular, 87.0% of the total publications are distributed according to the Matthew effect,[8] which results in the highly-skewed distributions of the institutions' publication records. On the contrary, accelerated growth of collaboration works to restore the homogeneity of the institutions.

The rest of the paper is organized as follows. Section 2 presents the definitions and empirical methods and results. Section 3 proposes an explanation of the empirical power laws, based on an accelerated network model. Section 4 concludes.

2. Power laws for financial research institutions

2.1 Definitions

A power law, also referred to as a scaling law or a Pareto law, for a random variable x is a distribution described by the density function

$$f(x) = \frac{\alpha}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\alpha-1}, \quad x \geq x_{\min}, \alpha > 0 , \tag{1}$$

where α is the power exponent and $x_{\min}$ is the lower bound to the power law behavior. To avoid divergence as $x \to 0$, we require $x_{\min} > 0$. The countercumulative distribution function is defined to be $\bar{F}(x) = \Pr(X > x)$. We have

$$\bar{F}(x) = \int_x^\infty f(z) dz = \left(\frac{x}{x_{\min}} \right)^{-\alpha} . \tag{2}$$

[8] Bol et al. (2018) document the Matthew effect in the allocation of science funding in the Innovation Research Incentives Scheme, a primary funding source for young Dutch scientists, in the Netherlands.

Power laws are an important empirical regularity for a wide spectrum of natural and social phenomena (Dorogovtsev and Mendes, 2003; Jackson, 2008). Beginning with Pareto (1896), well-documented power laws in economics and finance include individual wealth/income, international trade, city size, stock-market activity, firm size, CEO compensation, and supply of regulations (Gabaix, 2009). In academia, the number of papers written by individual scientists, the number of coauthors of mathematicians, and the number of citations of papers also obey power laws (Newman, 2003). In the following, we provide a first characterization of the distributions of top finance journal publications on the institution level. Power laws prove to adequately describe these distributions in the upper tail.

2.2 Data

We manually collect the author and affiliation information of all papers published in the top three finance journals by reading the title page of each paper. We exclude notes, comments/replies, and speeches/addresses. We focus on a sample since 1995, which embodies the fast growth of cross-affiliation in financial research as shown in Figure 1. Until 2016, we register 4796 papers associated with 10910 author counts and 13548 affiliation counts. Both counts are with repetitions, that is, each mentioning of an author or institution contributes one count to its own category.[9] The total affiliation counts are then sorted and ascribed to the 828 distinct institutions. We simply treat the counts as the institutions' numbers of publications because they naturally pick up the effect of cross-affiliation and give full scope to collaboration.[10]

[9] For a specific example, Edmans et al. (2012) supply the following acknowledgement of the af-b filiations of the authors on the title page of the *JF* article: "Edmans is from The Wharton School, University of Pennsylvania, NBER, and ECGI; Gabaix is from the NYU Stern School of Business, NBER, CEPR, and ECGI; Sadzik is from New York University; and Sannikov is from Princeton University." In our sample, this paper consists of four authors affiliated with six distinct institutions. And we count one publication for University of Pennsylvania, two for NBER (National Bureau of Economic Research), one for CEPR (Center for Economic and Policy Research), two for ECGI (European Corporate Governance Institute), two for New York University, and one for Princeton University.

[10] Similarly, Newman (2004), in the context of coauthorship networks, provides the simplest method of counting the frequency of coauthorships between pairs of individual scientists as a measure of the strength of collaboration.

2.3 Empirical methods

A power law model and the lower bound to the power law behavior can be determined adaptively. Given the cutoff for points in the upper tail, we get the measured data $x_1 \geq x_2 \geq \ldots \geq x_N = x_{min}$ with N observations. We estimate α as the slope of the OLS log-log regression (Gabaix and Ibragimov, 2011):

$$\ln(n - \gamma) = constant - \alpha \text{b} \ln x_n + noise, \tag{3}$$

where γ is an adjustment to the rank n. $\gamma = 0$ is typical in the literature while $\gamma = 1/2$ is optimal. The downward adjustment of 1/2 to the rank reduces the small-sample bias to the leading order for the log-log regression.[11] The asymptotic standard error of αb is computed at $\alpha\text{b} \cdot (N/2)^{-1/2}$, not the conventional OLS standard error that is nullified by the positive autocorrelation introduced by the ranking procedure into the residuals.

The lower bound now hinges on a tradeoff: Too high x_{min} will leave out legitimate data points thus increase finite sample biases, while too low x_{min} introduces biases by mixing with non-power-law data. We adopt the approach proposed by Clauset, Shalizi, and Newman (2009), which chooses x_{min} to minimize the distance between the empirical distribution of the data and the fitted power law. We measure the distance using the Kolmogorov-Smirnov statistic, D_{KS}, which is most common for nonmormal data:

[11] Barro and Jin (2011) impose the same downward adjustment prescribed by Gabaix and Ibragimov (2011) when estimating their power laws for the macroeconomic disasters.

$$D_{KS} = \max_{x \geq x_{min}} |P(x) - F(x)|, \tag{4}$$

where $P(x)$ and $F(x)$ are the CDFs of the data and the fitted power law model, respectively. Our method avoids the subjectivity in applying some upper quantile or visual checks to fix $x_{\min}$ as often done in the literature (see, e.g., Beirlant et al., 2004 and Drees, De Haan, and Resnick, 2000), while other more subtle methods require estimation of extra parameters (Embrechts, Kluppelberg, and Mikosch, 1997).

A two-parameter distribution, such as the lognormal, may provide a better fit for the data due to the added curvature from the free parameter (Eeckhout, 2004). To test potential deviations from a power law, we employ the Gabaix-Ibragimov test (Gabaix and Ibragimov, 2008) which augments the OLS regression with a quadratic term:

$$\ln(n - 1/2) = constant - \widehat{\alpha} \ln x_n + \widehat{\beta}(\ln x_n - \bar{x})^2 + noise, \tag{5}$$

where $\bar{x} \equiv cov((\ln x_n)^{-2}, \ln x_n)/var(\ln x_n)/2$. $\bar{x}$ recenters $\ln x_n$ such that αb stays theb same, irrespective of the quadratic term. The standard error of the test statistic β is computed at αb$^2 \cdot (2N)^{1/2}$. We reject the null of an exact power law if and only if βb is statistically different from zero.

2.3 Empirical findings

We show in Panel A of Figure 2 the distribution of top finance journal publications accumulated to the end of 2016 for all institutions. We plot the rank of the institutions against their total publications in log-log coordinates. We impose the optimal downward adjustment of 1/2 to the rank as suggested by Gabaix and Ibragimov (2011). The upper tail of the distribution falls onto a straight line that hints at a power law, or a rank-size rule, as dictated by (2).

Applying the method discussed in the previous subsection, we retain the 50 most prolific institutions in the upper tail, with the estimated cutoff of $x_{min} = 71$. We report the complete list of the 50 institutions together with their publication records in Table 1. Although the 50 institutions account for merely 6.0% of the entire sample of institutions that had published in top finance journals during our sample period, they contribute a predominant 61.3% of total publications. The distribution of the 50 institutions is highly right-skewed, with a mean number of publications of 166 and a median of 112; notably, NBER is solely responsible for 1105 publications.

We classify the 50 institutions according to several criteria in Table 2. The comprehensiveness of our sample is revealed by the scattering of the institutions across type, geographic region, and time of establishment. For instance, the majority of the most productive institutions, as expected, are universities and colleges. Meanwhile, we find that three research bureaus/centers (NBER, CEPR (Centre for Economic Policy Research), and ECGI (European Corporate Governance Institute)), the Federal Reserve Board, and the Federal Reserve Bank of New York also make their way into the upper tail of the distribution. It is clear from the table that the present study broadly includes all institutions that embark on academic research in financial economics as “research institutions”.

We present the distribution of the 50 most prolific institutions in Panel B of Figure 2. A power law model describes the data remarkably well (for over two decades of institutions and from 10^1 to 10^3 publications). We report the OLS regression results in Table 3. For the three journals combined, the regression R^2 is as high as 98.6%. The power exponent is estimated at 1.640 with a standard error of 0.328. As a robustness check, we repeat the analysis for the preceding five years, from 2011 to 2015. The results are presented in Table 4. The power exponent was fairly stable, even though the average number of publications increased dramatically over time. The complex system of institutions, with diverse natures and worldwide origins, seems to have reached a scale-free stationary state.

To explore the unanimity of the power law behavior, we redo the graph of the upper tail distribution for each journal separately in Panel C of Figure 2. Since the three journals publish different total numbers of papers, we normalize each institution’s publications in one journal by the average number of publications in that journal. If the same power law permeates all three journals, we expect their graphs to collapse onto one another after normalization. This is confirmed in Panel C of Figure 2. Moreover, Table 3 shows that, for each journal separately, the regression R^2 is no less than 96.1%, and the power exponent ranges from 1.490 to 1.689 with no statistically significant differences among those estimates given the standard errors.

We present in Table 5 the results of the Gabaix-Ibragimov test for deviations from our proposed power law models. For all the three journals, separately or combined, we see that βb is not significantly different from zero given the standard errors. We fail to significantly improve model performance by introducing curvature to the models.

For a final robustness check, we apply the well-known tail-index estimator of Hill (1975) to our empirical data. The estimator is

$$\widehat{\alpha}^H = \left(\frac{1}{N-1} \sum_{n=1}^{N-1} (\ln x_n - \ln x_N) \right)^{-1}, \tag{6}$$

with a standard error of $\alpha b^H \cdot N^{-1/2}$.[12] This estimator has been reported in, e.g., Dobkins and Ioannides (2000). We use the Kolmogorov-Smirnov test of the empirical CDF against the Hill's estimate to detect potential deviations from power laws. We present the results for model estimation in Panel A and model tests in Panel B of Table 6. We confirm that, although some variations in the results are unavoidable when we apply different estimation methods, we find no qualitative differences in the results under either the OLS log-log regression or the Hill's tail-index estimator approach. Specifically, the discrepancies in the power exponent estimates are not statistically significant given the standard errors for the power law models considered. Furthermore, the Kolmogorov-Smirnov test cannot reject the power law models at a conventional significance level.

In sum, the proposed power laws seem to succinctly capture the essence of the growth in financial research institutions. We now turn to examine the economic forces behind and dynamic evolutions of the academic finance research society, which may have effectively engendered the scaling laws documented above.

[12] Under the null of a perfect power law, the Hill's estimator exploits the efficiency properties of maximum likelihood estimation, and hence provides a smaller standard error than that of Gabaix and Ibragimov (2011).

3. A network-based explanation

A variety of processes can explain the emergence of scaling (see, e.g., Barab´asi and Albert, 1999; Gabaix, 1999; and Ferrer i Cancho and Sol´e, 2003).[13] We find empirical support for the network model of Barab´asi and Albert (1999), who pioneer two generic mechanisms: One is *growth*, that is, new nodes continue entering the network; the other is *preferential attachment*, also referred to as *success-breeds-success*, *rich-get-richer*, *popularity-is-attractive*, or *Matthew effect*, that is, a node with a higher degree will enjoy a cumulative advantage and acquire even more links in the future.

Specifically, we treat institutions with at least one top finance journal publication as nodes in our collaboration network. Therefore, an institution joins the network when it publishes for the first time in the top finance journals. And we treat the total number of publications of a node as the degree of the node. In concert with our data acquisition method, we increase the degree of a node by one per citation of the corresponding affiliation by an author in an article. We plot in Figure 3 the structure of the collaboration network at the end of 2016. Two nodes are connected if the corresponding two institutions work together in at least one publication during 19952016. Notably, the network has a giant component that includes 99.5% of the nodes, which reveals the close linkage among the academic society for financial research.

3.1 Dynamic patterns

We first examine the growth of the collaboration network of financial research institutions. We plot in Panel A of Figure 4 the number of nodes newly joining the network. We see that the growth of the nodes is stable, with certain fluctuations around the financial crisis of 2008. Panel B of Figure 4 presents the fraction of publications contributed by the new nodes for each year from 1995 to 2016. This proportion again stays relatively stable, with an average of 6.0% from 2000 to 2016.

We then investigate the allocation of new links (publications) to the nodes. Exact power laws follow only from *linear* preferential attachment (Krapivsky, Redner, and Leyvraz, 2000). That is, the probability for an existing node to acquire new links ($\Pi(k)$) is proportional to its

[13] For textbook treatments on the topic, see Dorogovtsev and Mendes (2003), Jackson (2008), Baraba´si (2016), and Newman (2018), among others.

degree (k), or $\Pi(k) \sim k$. Exploiting the network maps at the end of each year from 1995 to 2016, we use the method provided in Jeong, N´eda, and Barab´asi (2003) to measure preferential attachment. We provide detailed results in Appendix A. We find that $\Pi(k)$ can be well approximated by $\Pi(k) \sim k^{\theta}$, where θ is estimated at 1.00 with a standard error of 0.08. Thus, linear preferential attachment achieves a good match for data.

In the following, we develop a network model that captures three additional features of the collaboration network. First, expanding coauthorships (Panel A of Figure 1) and cross-affiliation (Panel B of Figure 1) generate nonlinear growth in total publications. Thus, instead of linear growth considered by Barab´asi and Albert (1999), we incorporate accelerated growth following Dorogovtsev and Mendes (2001b, 2003). Second, we include internal links because 94.0% of the total publications come from collaboration among already existing nodes (Ghoshal, Chi, and Barab´asi, 2013). Third, although preferential attachment enables institutions successful in historic records to produce even more research output in the future, we randomly allocate a fraction of new publications to the existing nodes, allowing research opportunities to strike institutions by sheer chance (Ghoshal, Chi, and Barab´asi, 2013).[14]

3.2 An evolving network with accelerated growth

Consider discrete time $t = 0,1,2,...,$ with time interval Δt. Assume that there is one node in the network at time 0 and that one new node enters the network at each $t \geq 1$. Then there are t nodes in the network at time t, excluding the new node entering at the time. Let $m(t)$ denote the publication rate at time t. Hence $m(t)\Delta t$ is the total number of new publications during the time interval $[t,t + 1]$. These publications are distributed among all nodes in the network at time $t + 1$. We assume that, a faction c_0 of the total publications is distributed *preferentially* to the existing nodes, a fraction c_1 is distributed *randomly* (with equal probability $1/t$) to the existing nodes, and a fraction c_2 is allocated to the new node. The constants c_0, c_1, and c_2 are nonnegative and sum to one.[15] Preferential and random allocations are independent events.

Let $X_{s,t}$ denote the degree at time t of a node entering at time s. Notice that

[14] Mele (2017) develops a structural model of network formation, with both strategic and random networks features, which generates directed dense networks. He also provides a Bayesian MCMC method to estimate model parameters.

[15] For simplicity, we assume $m(0)\Delta t$ goes to the only node at $t = 0$ in the network.

Pr(Xs,t+1 = k) = XPr(Xs,t+1 = $k,Xs,t = j$)

j≥0

= $\Pr(X_{s,t+1} = k|X_{s,t} = k - (c_0 + c_1)m(t)\Delta t)\Pr(X_{s,t} = k - (c_0 + c_1)m(t)\Delta t)$

+ $\Pr(X_{s,t+1} = k|X_{s,t} = k - c_1m(t)\Delta t)\Pr(X_{s,t} = k - c_1m(t)\Delta t)$ (7)

+ $\Pr(X_{s,t+1} = k|X_{s,t} = k - c_0m(t)\Delta t)\Pr(X_{s,t} = k - c_0m(t)\Delta t)$ + Pr(Xs,t+1 =

$k|Xs,t = k$)Pr($Xs,t = k$).

The four conditional probabilities in the above equation can be evaluated by linear preferential attachment, random allocation, and independence of the events.

Let $q(k,s,t) = \Pr(X_{s,t} = k)$ denote the probability that the node has degree k at time t, $t \geq s$. A new node entering at time s obtains the degree $c_2m(s - 1)\Delta t$. Thus, $q(c_2m(s-1)\Delta t,s,s) = 1$. We have the following master equation for degree probabilities of individual nodes.

$$\begin{aligned} q(k,s,t+1) = {} & \frac{k-(c_0+c_1)m(t)\Delta t}{A(t)}\frac{1}{t}q(k-(c_0+c_1)m(t)\Delta t,s,t) \\ & + \left(1-\frac{k-c_1m(t)\Delta t}{A(t)}\right)\frac{1}{t}q(k-c_1m(t)\Delta t,s,t) \\ & + \frac{k-c_0m(t)\Delta t}{A(t)}\left(1-\frac{1}{t}\right)q(k-c_0m(t)\Delta t,s,t) \\ & + \left(1-\frac{k}{A(t)}\right)\left(1-\frac{1}{t}\right)q(k,s,t), \quad s \leq t,\ t=1,2,..., \end{aligned} \tag{8}$$

where $A(t) = \sum_{s=0}^{t-1} m(s)\Delta t$ is the cumulative degrees of all nodes at t. It is worth mentioning that $k \leq A(t)$, that is, the degree of a node at time t is not greater than the cumulative input of degrees until time t. Re-arrange the terms of (8), we have:

$$\begin{aligned} \frac{q(k,s,t+1)-q(k,s,t)}{\Delta t} = {} & \frac{1}{tA(t)\Delta t}((k-(c_0+c_1)m(t)\Delta t)q(k-(c_0+c_1)m(t)\Delta t,s,t) \\ & -(k-c_1m(t)\Delta t)q(k-c_1m(t)\Delta t,s,t)) \\ & + \frac{1-1/t}{A(t)\Delta t}((k-c_0m(t)\Delta t)q(k-c_0m(t)\Delta t,s,t)-kq(k,s,t)) \\ & + \frac{1}{t\Delta t}(q(k-c_1m(t)\Delta t,s,t)-q(k,s,t)). \end{aligned}$$

To exploit the tractability of the rate-equation approach (Dorogovtsev, Mendes, and Samukhin, 2000; Krapivsky and Redner, 2002), we let Δt go to zero as time t and degree k vary continuously. Suppose the degree rate function $m(t)$ is continuous. The continuous-time approximation of the master equation (8) is obtained as follows. For $t > s \geq 0$,

$$\frac{\partial p(k,s,t)}{\partial t} + \frac{c_0 m(t)}{A(t)}\frac{\partial kp(k,s,t)}{\partial k} + \frac{c_1 m(t)}{t}\frac{\partial p(k,s,t)}{\partial k} = 0, \tag{9}$$

where $p(k,s,t)$ is a generalized probability density function representing the density of the degree distribution at time t of the node s. The initial condition is $p(k,s,s) = \delta(k - c_2 m(s))$ for $s >$ 0, where $\delta(\cdot)$ is the delta function with the properties $\delta(0) = \infty$, $\delta(x) = 0$ for x 6= 0, and $\int_0^\infty \delta(k-k_0)dk = 1$ for a non-negative constant k_0.

Multiplying by k and integrating from 0 to ∞ with respect to k for the above equation, we obtain:

$$\frac{\partial \int_0^\infty kp(k,s,t)dk}{\partial t} = \frac{c_1 m(t)}{t} + \frac{c_0 m(t) \int_0^\infty kp(k,s,t)dk}{A(t)}. \tag{10}$$

Denote the average degree $\bar{k}(s,t) = \int_0^\infty kp(k,s,t)dk$. Then we obtain the evolution equation for $\bar{k}(s,t)$:

$$\frac{\partial \bar{k}(s,t)}{\partial t} = \frac{c_1 m(t)}{t} + \frac{c_0 m(t)\bar{k}(s,t)}{A(t)}, \tag{11}$$

with boundary condition $\bar{k}(s,s) = \int_0^\infty kp(k,s,s)dk = c_2 m(s).$

The integration $\int_0^t p(k,s,t)ds$ is the total number of nodes with degree k at time t, and t is the number of nodes at time t, excluding the newly coming one. Hence the degree distribution at time t is given by $P(k,t) = \frac{1}{t}\int_0^t p(k,s,t)ds$. Actually, the distribution $P(k,t)$ can be recovered from $\bar{k}(s,t)$ by the relation

$$P(k,t) = \frac{1}{t}\int_0^t \delta(k - \bar{k}(s,t))ds = -\frac{1}{t}\left(\frac{\partial \bar{k}(s,t)}{\partial s}\right)^{-1}\Bigg|_{s=\hat{s}(k,t)}, \tag{12}$$

where $s = \hat{s}(k,t)$ is a solution to the equation $k = \bar{k}(s,t)$. The first equality is due to the fact that the solution $p(k,s,t)$ is actually a delta function $p(k,s,t) = \delta(k - k_s)$ for some k_s and the second equality is according to the general property of the delta function.

We now turn to solve the equation (11). A general solution for $\bar{k}(s,t)$ is

$$\bar{k}(s,t) = \left(C(s) + \int_1^t \frac{c_1 m(v)}{v} \left(\int_0^v m(u)du \right)^{-c_0} dv \right) \left(\int_0^t m(v)dv \right)^{c_0}. \quad (13)$$

By the boundary condition $\bar{k}(s,s) = c_2 m(s)$, $C(s)$ can be identified and we obtain

$$\bar{k}(s,t) = c_2 m(s) \left(\frac{\int_0^s m(v)dv}{\int_0^t m(v)dv} \right)^{-c_0} + \int_s^t \frac{c_1 m(v)}{v} \left(\frac{\int_0^v m(u)du}{\int_0^t m(u)du} \right)^{-c_0} dv. \quad (14)$$

Recall that $\bar{k}(s,t)$ is the average degree at time t of the nodes entering at time s. The above equation indicates that a lower s corresponds to a higher $\bar{k}(s,t)$ for a fixed t.

Let $s = \hat{s}(k,t)$ be a solution to the equation $\bar{k}(s,t) = k$. Then $\bar{k}(\hat{s}(k,t),t) = k$. Taking partial derivative with respect to k, it follows that

$$\frac{\partial \bar{k}}{\partial k} = \left. \frac{\partial \bar{k}}{\partial s} \right|_{s=\hat{s}} \frac{\partial \hat{s}}{\partial k} = 1 . \quad (15)$$

Hence

$$\left. \frac{\partial \bar{k}}{\partial s} \right|_{s=\hat{s}} = \left(\frac{\hat{s}}{\partial k} \right)^{-1},$$

where $\hat{s} = \hat{s}(k,t)$ satisfies

$$k = c_2 m(\hat{s}) \left(\frac{\int_0^{\hat{s}} m(v)dv}{\int_0^t m(v)dv} \right)^{-c_0} + \int_{\hat{s}}^t \frac{c_1 m(v)}{v} \left(\frac{\int_0^v m(u)du}{\int_0^t m(u)du} \right)^{-c_0} dv. \quad (16)$$

Next we introduce an assumption on the rate function $m(t)$ in order to evaluate $\partial \bar{k}(s,t)/\partial s$. Suppose that $m(t)$ is a non-negative continuous function defined on $[0,\infty)$.

Assumption: There exist constants $\gamma \geq 0$, $b \geq a > 0$, such that $as^\gamma \leq m(s) \leq bs^\gamma$ for all $s > 0$.

For example, $m(s) = m$ or $m(s) = ms^\gamma$ for some positive constant m, satisfies the assumption. Clearly, when the assumption is satisfied, $m(s)$ is such a function that $\limsup_{s\to\infty} \frac{m(s)}{s^\gamma}$, $\liminf_{s\to\infty} \frac{m(s)}{s^\gamma}$, $\limsup_{s\to 0^+} \frac{m(s)}{s^\gamma}$, and $\liminf_{s\to 0^+} \frac{m(s)}{s^\gamma}$ all exist. We write

$h(s) = O(s^{\gamma})$ if a continuous function $h(s)$ satisfies the assumption. We can directly show that $\int_0^s h(\nu)d\nu = O(s^{\gamma+1})$ if $h(s) = O(s^{\gamma})$.

Using the above assumption, we can estimate the right hand side of (14). We find that

$$k = f(t)O(\hat{s})^{\gamma-c_0(1+\gamma)} + g(t), \tag{17}$$

where

$$f(t) = (c_2 - \frac{c_1}{\gamma - c_0(1+\gamma)})\left(\int_0^t m(v)dv\right)^{c_0} = O(t)^{c_0(1+\gamma)},$$
$$g(t) = \frac{c_1}{\gamma - c_0(1+\gamma)}\left(\int_0^t m(v)dv\right)^{c_0} O(t^{\gamma-c_0(1+\gamma)}) = O(t)^{\gamma}.$$

Thus, $\hat{s} = (f(t))^{-1}O(k-g(t))^{\frac{1}{\gamma-c_0(1+\gamma)}}$ and

$$\frac{\partial \hat{s}}{\partial k} = \frac{(f(t))^{-1}O(k-g(t))^{-\left(1-\frac{1}{\gamma-c_0(1+\gamma)}\right)}}{\gamma - c_0(1+\gamma)} = \frac{f(t)^{-1}O(k)^{-\left(1-\frac{1}{\gamma-c_0(1+\gamma)}\right)}}{\gamma - c_0(1+\gamma)}. \tag{18}$$

The second equality in the above equation holds when k is large and has the same or a higher order than $O(t)^{\gamma}$ of $g(t)$. Note that the total degree $\int_0^t m(v)dv$ has an order of $O(t)^{1+\gamma}$ thus the average degree $\frac{1}{t}\int_0^t m(v)dv$ has an order of $O(t)^{\gamma}$. Finally, it follows from (15) and (12) that the degree distribution at time t is

$$P(k,t) = t^{-1}f(t)^{-1}O(k)^{-\left(1+\frac{1}{c_0(1+\gamma)-\gamma}\right)}, \tag{19}$$

or

$$P(k,t) \sim k^{-\left(1+\frac{1}{c_0(1+\gamma)-\gamma}\right)}, \tag{20}$$

when k is large. Therefore, the degree distribution obeys a power law in the upper tail with exponent

$$\alpha = \frac{1}{c_0(1+\gamma)-\gamma}. \tag{21}$$

To ensure the existence of k^{-}, we require 1 +$\alpha > 2 \text{ or } 1 > c_0 > \frac{\gamma}{1+\gamma}$.

Remark 1 (Additional Attractiveness): *Dorogovtsev and Mendes (2001b) first study scaling property of networks with accelerated growth, based on the model of Albert and Barabasi (1999). To obtain a power exponent (α) greater than 1, they introduce a timedependent "additional attractiveness" to the nodes, which may not be easily verified empirically (Jeong, N´eda, and Barab´asi, 2003). Our model achieves the scaling property of networks with nonlinear growth without this extra assumption.*

Remark 2 (Experience versus Talent): *Kong, Sarshar, and Roychowdhury (2008) consider the degree of a node as a proxy of its experience, and provide a method to identify the talent (inherent fitness) of a node. For instance, the talent of an institution in our study may be the endowment, scale, location, pedigrees of the employees, etc., of the institution. All these factors may potentially determine the institution's future research performance. Under the combined effects of experience and talent, new degrees are then allocated preferentially according to the product of experience and talent.*[16] *Kong, Sarshar, and Roychowdhury (2008) find that the talents of webpages are exponentially distributed, which explain the power law degree distributions of webpages and the quick rising of interesting new pages in the page ranking. However, the same method does not generate significant differences in the talents of the research institutions in our study.*[12] *We thus do not pursue extensions of our model along this dimension.*

3.3 Model implications

From (21), the exponent α decreases in c_0. The larger proportion of degree allocation according to preferential attachment heightens the success-breeds-success effect and enhances the dispersion of the degree distribution. Suppose that c_0 is strictly less than 1. Then α increases in γ. Heterogeneity of the system diminishes as the network undergoes higher order growth. Intuitively, fast growth of the network brings more and more publications for random allocation and allocation to new nodes, the two channels that counterbalance the effect of preferential attachment and restore homogeneity of the system. Moreover, preferential attachment works by cumulative advantages that are tempered by more and more publications

[16] Note that the effect of "additional attractiveness" diminishes as the degree of a node becomes

to be allocated in the futures. For example, after a period of fast growth of the network, a newly joining node may have a degree higher than that of an existing node that has successfully received several preferential allocations. At the limit of $\gamma \to c_0/(1 - c_0)$, the power law collapses and heterogeneity of the nodes disappears. c_1 and c_2 do not directly enter (21), but they indirectly affect α through c_0 as $c_0+c_1+c_2 = 1$. For $c_0 \to 1^-$, we get $\alpha \to 1^+$, a special case called Zipf's law, which is interesting because all its moments diverge. γ does not play a role in this case. We obtain the Barab´asi and Albert model in the special case of $c_0 = c_2 = 0.5$ and $\gamma = 0$.

For our collaboration network, the total counts of new publications equal the product of the number of papers, the average number of authors per paper, and the average number of affiliations per author. Coauthorships and cross-affiliation exhibit approximately linear growth, suggesting the quadratic growth of $m(t)$ with $\gamma = 2$, given that the total number of papers published annually remains stable in the long run. Our empirical study finds that αb = 1.64 for the three journals altogether. We can use (21) to obtain c_0 = 87.0%. That is, 87.0% of publications may have resulted from success-breeds-success, a dominating channel for the distribution of the top journal publications. We have deduced from the data that 6.0% of publications go to new nodes. Random allocation thus takes the remaining 7.0% of publications. When we lack either the growth of coauthorships or the growth of cross-affiliation, γ is brought down to 1, which

larger, while the effect of talent does not.

[12]The results are available upon request.

implies c_0 = 80.5% under slowed acceleration; additionally, when we lack both forms of collaboration, we have $\gamma = 0$ and c_0 = 61.0% under linear growth. This comparison emphasizes the importance of identifying the collaboration patterns before further inference on the evolving network.

A most commonly invoked measure of dispersion/inequality in studies of income or wealth distributions is the Gini coefficient (G). $G = 0$ represents perfect equality while $G = 1$ stands for the other extreme of maximal inequality. Dispersion increases with G monotonically in between. $G = 0.5$ is commonly regarded as the warning level for great disparity. We here borrow this measure and provide further discussions on the heterogeneity of the institutions. For power law distributions, it is well-known that $G = 1/(2\alpha - 1)$ (Aaberge, 2005). For the three journals altogether, we have $G = 0.44$ for αb = 1.64. Given $c_0 = 87.0\%$, $c_1 = 7.0\%$, and $c_2 = 6.0\%$, when either the growth of coauthorships or the growth of cross-affiliation is absent, α reduces to 1.35 implying $G = 0.59$. And again, given the shares, when both forms of collaboration are absent, α reduces to 1.15 implying $G = 0.77$. Thus, higher order growth in collaboration effectively diminishes heterogeneity in the institutions' research output.

4. Conclusion

Institutions are fundamental units for conducting research activities. Cross-affiliation, nonlinear growth from mixed forms of collaboration, and power laws for institutions' research output may characterize other fields and extend to a broader set of journals as well. Cross-field examinations of the scaling behavior and the role played by preferential attachment in the allocation of publications could undoubtedly enrich our understanding of production of scientific knowledge on the institution level. We can also examine the inverse relationship between order of growth in collaboration and heterogeneity of research institutions through cross-disciplinary studies. This work awaits comprehensive data to be gathered for collaboration networks of different disciplines.

Appendix A. Measuring preferential attachment

At every time step t, a node i already present in the network acquires new links at the rate $\Pi(k_i)$ where k_i is the degree of the node. Under preferential attachment, $\Pi(k_i)$ is a monotonically increasing function of k_i. Several authors propose that Π follows a power law (Krapivsky, Redner, and Leyvraz, 2000; Newman, 2001a; Jeong, N´eda, and Barab´asi, 2003):

$$\Pi(k_i) = \frac{k_i^{\theta}}{\sum_j k_j^{\theta}} = C_t k_i^{\theta}, \tag{A.1}$$

where $\theta > 0$ is a scaling constant; C_t is a normalization pertinent to time t. The scalefree model of Barab´asi and Albert (1999) corresponds to linear preferential attachment with $\theta = 1$, while for $\theta < 1$ (sub-linear), the degree distribution follows a stretched exponential and for $\theta > 1$ (super-linear), a single node connects to nearly all other nodes (Krapivsky, Redner, and Leyvraz, 2000). We use the method provided in Jeong, N´eda, and Barab´asi (2003) to examine whether Π could be well approximated by a power law, and if so, to estimate the exponent θ.

The dynamics of the research institution network provide the time when each node joins the system and the degrees of the nodes from 1995 to 2016. If the evolving network develops a stationary state, we can use nodes already present at time T and the degrees of these nodes at both T and T +ΔT to measure Π. To ensure stationarity, we need T away from T0 when the network starts. We also choose a small ΔT such that the effect of t is at minimum and Π relies exclusively on k. To further reduce noise, we examine the cumulative function

$$\pi(k) = \sum_{k_i=1}^{k} \Pi(k_i). \tag{A.2}$$

For $\Pi(k) \sim k^{\theta}$, we expect $\pi(k) \sim k^{\theta+1}$.

In our empirical implementation, we choose $T_0 = 1995$, T =2000-15, and $\Delta T = 1$. The $\pi(k)$ functions are shown in Panel A of Figure A.1. We plot two additional benchmark cases for comparison: $\pi(k) \sim k$ without preferential attachment and $\pi(k) \sim k^2$ for linear preferential attachment. The $\pi(k)$ functions of the network follow a straight line in log-log coordinates. Therefore, the power law form in (4) provides a good approximation. The $\pi(k)$ functions also increase with a speed consistent with linear preferential attachment, independently of the year when the measurements are taken. This result supports stationarity in the degree allocation process, even though some variation is inevitable due to statistical noise. We show the estimated θ exponent in Panel B of Figure A.1. The exponent from 2000-15 has a mean of 2.00 and a standard deviation of 0.08. Linear preferential attachment hence constitutes the key degree allocation mechanism of the growing network.

References

Aaberge, Rolf, “Gini’s Nuclear Family.” Journal of Economic Inequality, 5 (2007), 305-322.

Barab´asi, Albert-L´aszlo´, Network Science (Cambridge: Cambridge University Press, 2016).

Barab´asi, Albert-L´aszlo´, and R´eka Albert, “Emergence of Scaling in Random Networks,” Science, 286 (1999), 509-512.

Barab´asi, Albert-L´aszlo´, Hawoong Jeong, Zolt´an N´eda, Erzs´ebet Ravasz, Andr´as Schubert, and Tamas Vicsek, “Evolution of the Social Network of Scientific Collaborations,” Physica A, 311 (2002), 590-614.

Barro, Robert J., and Tao Jin, “On the Size Distribution of Macroeconomic Disasters,” Econometrica 79 (2011), 1567-1589.

Beirlant, Jan, Yuri Goegebeur, Johan Segers, and Jozef Teugels, Statistics of Extremes Theory and Applications, (Chichester: Wiley, 2004).

Bol, Thijs, Mathijs de Vaan, and Arnout van de Rijt, “The Matthew Effect in Science

Funding,” Proceedings of the National Academy of Sciences, 115 (2018), 48874890.

Card, David, and Stefano DellaVigna, “Nine Facts about Top Journals in Economics,” Journal of Economic Literature, 51 (2013), 144-161.

Clauset, Aaron, Cosma R. Shalizi, and Mark E.J. Newman, “Power-law Distributions in Empirical Data,” SIAM Review, 51 (2009), 661-703.

Overman, Henry G., and Yannis M. Ioannides, “Cross-Sectional Evolution of the U.S.

City Size Distribution,” Journal of Urban Economics, 49 (2001), 543-566.

Dorogovtsev, Sergey N., and Jos´e F.F. Mendes, “Effect of the Accelerating Growth of Communications Networks on Their Structure,” Physical Review E, 63 (2001a), 025101.

———, “Scaling Properties of Scale-free Evolving Networks: Continuous Approach,” Physical Review E, 63 (2001b), 056125.

———, Evolution of Networks: From Biological Nets to the Internet and WWW (New York: Oxford University Press, 2003).

Dorogovtsev, Sergey N., Jos´e F.F. Mendes, and Alexander N. Samukhin, “Structure of Growing Networks with Preferential Linking,” Physical Review Letters, 85 (2000), 4633.

Drees, Holger, Laurens De Haan, and Sidney Resnick, “How to Make a Hill Plot,” Annals of Statistics 28 (2000), 254-274.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov, “Dynamic CEO Compensation,” Journal of Finance, 67 (2012), 1603-1647.

Eeckhout, Jan, “Gibrat’s Law for (All) Cities,” American Economic Review, 94 (2004), 1429-1451.

Embrechts, Paul, Claudia Kluppelberg, and Thomas Mikosch, Modelling Extremal Events for Insurance and Finance, (New York: Springer, 1997).

Ferrer i Cancho, Ramon, and Ricard V. Sol´e, Optimization in Complex Networks In: Pastor-Satorras R., Rubi M., Diaz-Guilera A. (eds) Statistical Mechanics of Complex Networks. Lecture Notes in Physics, vol 625 (2003), pp.

114-125, Springer, Berlin.

Gabaix, Xavier, “Zipf’s Law for Cities: An Explanation,” Quarterly Journal of Economics, 114 (1999), 739-767.

———, “Power Laws in Economics and Finance,” Annual Review of Economics, 1

(2009), 255-293.

Gabaix, Xavier, and Rustam Ibragimov, “A Simple OLS Test of Power Law Behavior.” Working Paper, (2008), Harvard University.

———, “Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents,” Journal of Business Economics and Statistics, 29 (2011), 24-39.

Gagen, Michael J., and John Mattick, “Accelerating, Hyperaccelerating, and Decelerating Networks,” Physical Review E, 72 (2005), 016123.

Goyal, Sanjeev, Macro J. van der Leij, and Jos´e Luis Moraga-Gonz´alez, “Economics:

An Emerging Small World,” Journal of Political Economy, 14 (2006), 403-412.

Ghoshal, Gourab, Liping Chi, and Albert-L´aszlo´ Barab´asi, “Uncovering the Role of Elementary Processes in Network Evolution,” Scientific Reports, 3 (2013), 2920.

Hill, Bruce M., “A Simple Approach to Inference about the Tail of a Distribution,” Annals of Statistics, 3 (1975), 1163–1174.

Jackson, Matthew O., Social and Economic Networks (Princeton: Princeton University Press, 2008).

Jeong, Hawoong, Zolt´an N´eda, and Albert-L´aszlo´ Barab´asi, “Measuring Preferential Attachment in Evolving Networks,” Europhysics Letters, 61 (2003), 567-572.

Krapivsky, Paul L., and Sidney Redner, “Statistics of Changes in Lead Node in Connectivity-Driven Networks,” Physical Review Letters, 89 (2002), 258703.

Krapivsky, Paul L., Sidney Redner, and Francois Leyvraz, “Connectivity of Growing Random Networks,” Physical Review Letters, 85 (2000), 4629.

Kong, Joseph S., Nima Sarshar, and Vwani P. Roychowdhury, “Experience versus Talent Shapes the Structure of the Web,” Proceedings of the National Academy of Sciences, 105 (2008), 13724–13729.

Mele, Angelo, “A Structural Model of Dense Network Formation,” Econometrica, 85 (2017), 825-850.

Newman, Mark E.J., “Clustering and Preferential Attachment in Growing Networks,” Physical Review E, 64 (2001a), 025102.

———, “Scientific Collaboration Networks: I. Network Construction and Fundamental Results,” Physical Review E, 64 (2001b), 016131.

———, “Scientific Collaboration Networks: II. Shortest Paths, Weighted Networks, and Centrality,” Physical Review E, 64 (2001c), 016132.

———, “The Structure of Scientific Collaboration Networks,” Proceedings of the National Academy of Science, 98 (2001d), 404-409.

———, "The Structure and Function of Complex Networks," SIAM Review, 45 (2003), 167-256.

———, "Coauthorship Networks and Patterns of Scientific Collaboration," Proceedings of the National Academy of Science, 101 (2004), 5200–5205.

———, Networks, Second Edition (Oxford: Oxford University Press, 2018).

Pareto, Vilfredo, Cours D'Economie Politique (Geneva: Droz, 1896).

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi, "The Increasing Dominance of Teams in Production of Knowledge," Science, 316 (2007), 1036-1039.

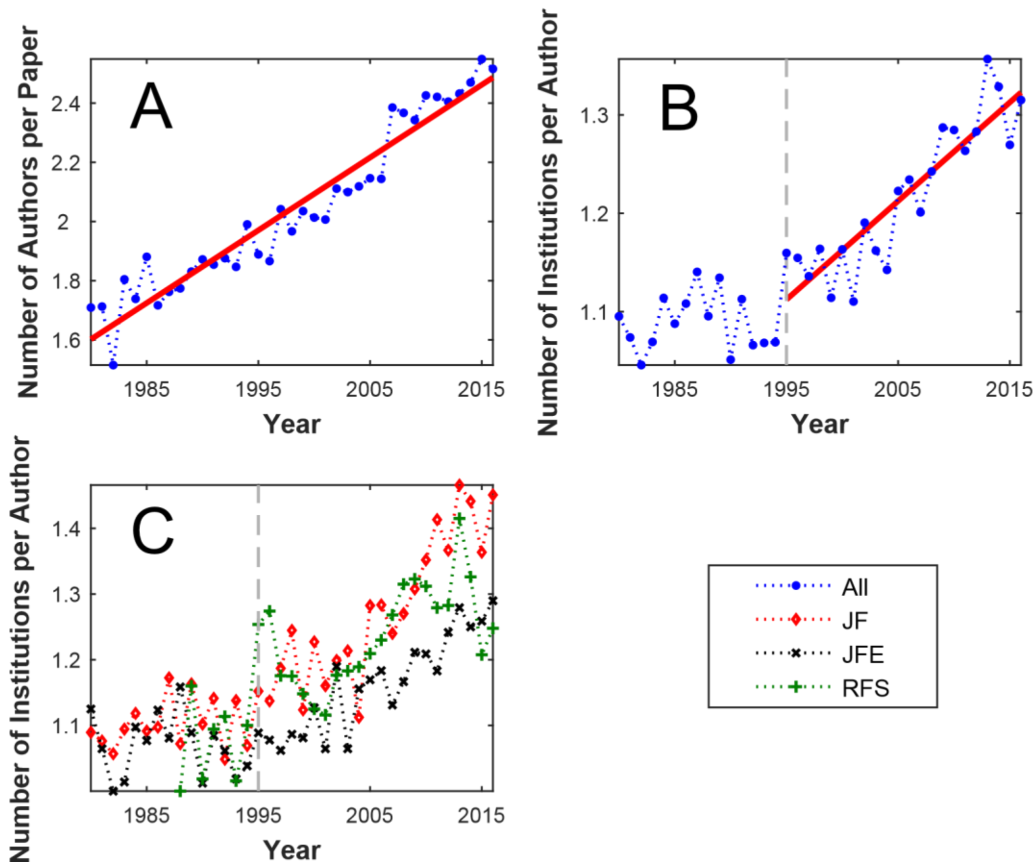


Figure 1. Coauthorships and cross-affiliations. We graph the average number of authors reported per paper ($\bar{N}_{aut}$) and the average number of institutions reported per author ($\bar{N}_{ins}$) at top finance journals (JF, JFE, and RFS). Each year, we calculate $\bar{N}_{aut}$ and $\bar{N}_{ins}$ using total paper, author, and affiliation counts in that year. (**A**) $\bar{N}_{aut}$ from aggregating the top three journals. (**B**) $\bar{N}_{ins}$ from aggregating the top three journals. (**C**) $\bar{N}_{ins}$ separately for each journal. The first issue of RFS appeared in 1988. The solid lines represent linear fit and vertical lines mark the year 1995.

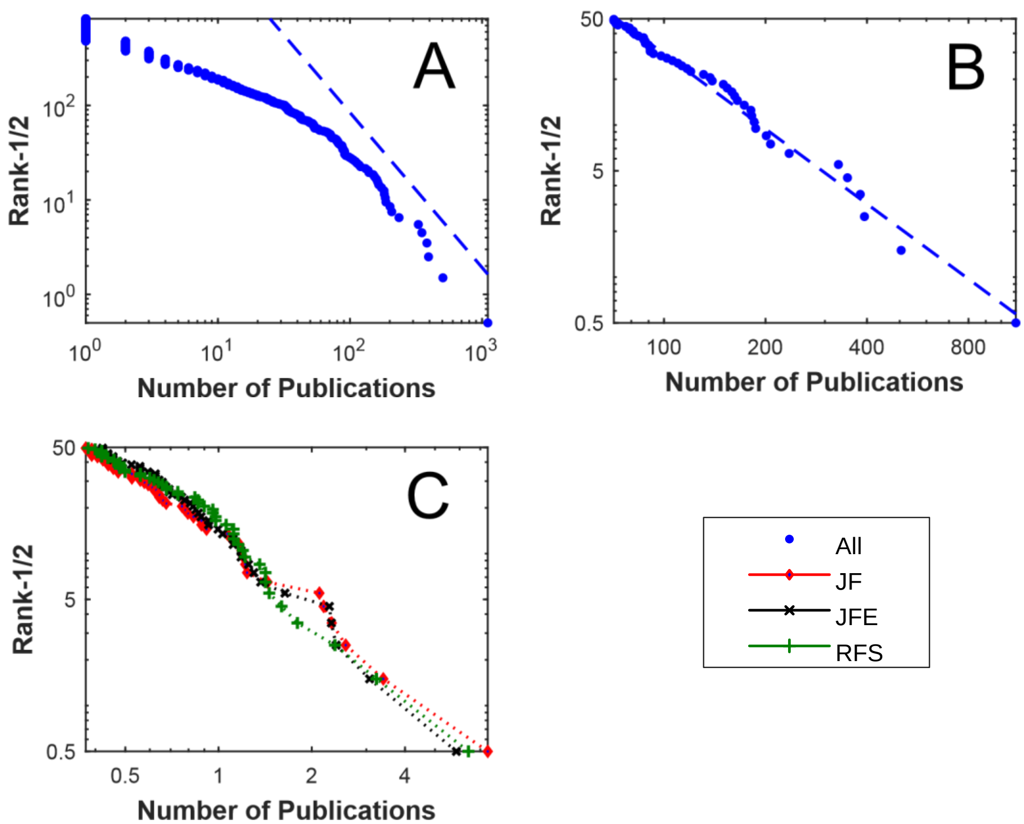


Figure 2. The distribution function of institutions' top finance journal publications. (**A**) Total publications in the top three journals (JF, JFE, and RFS) for all 828 institutions. (**B**) Total publications in the top three journals for the 50 most prolific institutions. (**C**) Total publications in each of the JF, JFE, and RFS for the 50 most prolific institutions.

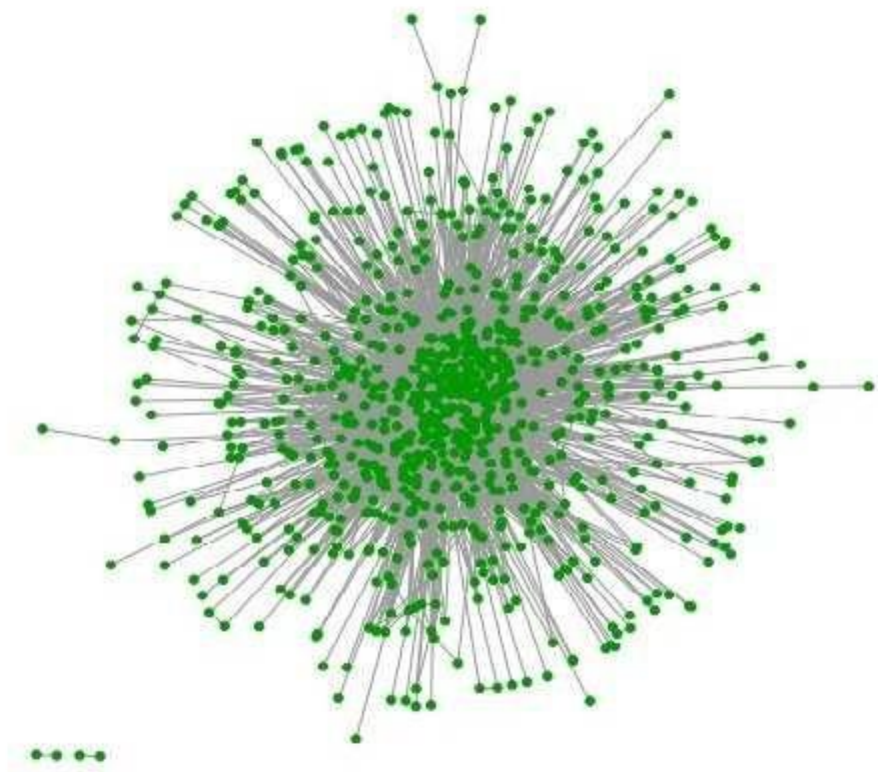
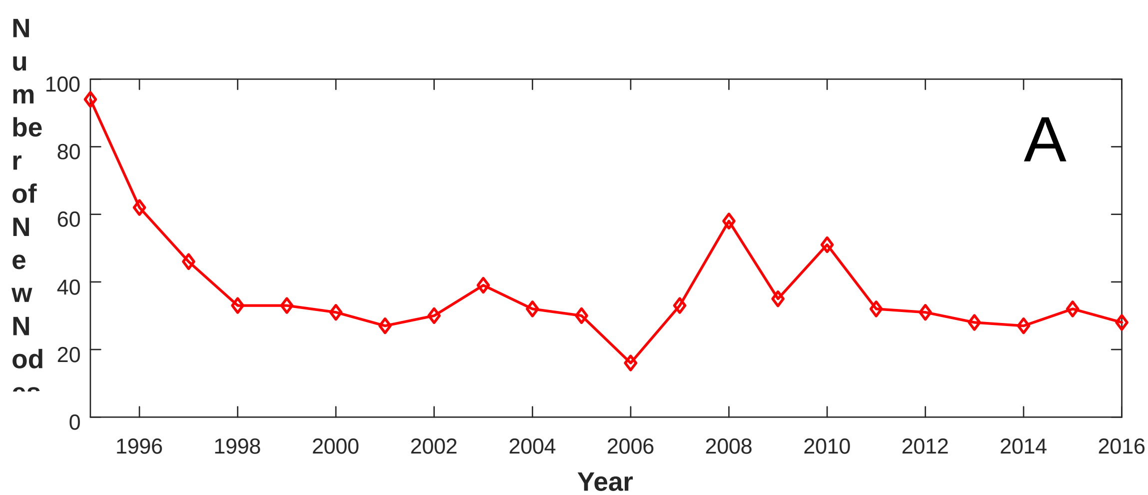


Figure 3. Collaboration network of financial research institutions. We link two institutions if they are simultaneously acknowledged, either by one author or different authors, in one publication from 1995 to 2016.



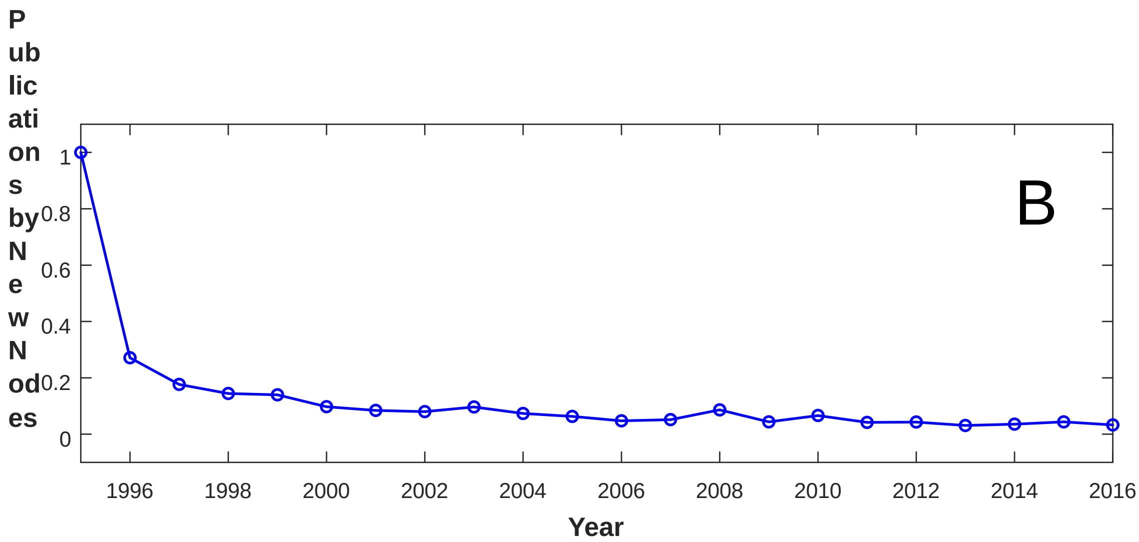


Figure 4. Growth of the collaboration network. (**A**) The number of newlyjoining nodes and (**B**) the proportion of publications by new nodes, for each studied year from 1995 to 2016.

Table 1. Top Finance Journal Publications by Institutions. This table reports the journal publications of the top 50 institutions from 1995 to 2016. The rank is based on aggregate publications across the three journals.

Rank	Institutions	JF	JFE	RFS	All Journals	% of Total
1	NBER	437	347	321	1105	13.30%
2	CEPR	201	142	162	505	6.08%
3	New York University	137	137	119	393	4.73%
4	Harvard University	129	182	71	382	4.60%
5	University of Pennsylvania	125	135	90	350	4.21%

6	University of Chicago	152	97	80	329	3.96%
7	Columbia University	85	77	73	235	2.83%
8	UCLA	70	81	56	207	2.49%
9	Duke University	73	70	58	201	2.42%
10	Stanford University	65	66	56	187	2.25%
11	MIT	72	70	43	185	2.23%
12	University of Michigan	52	59	71	182	2.19%
13	Ohio State University	47	74	60	181	2.18%
14	London Business School	54	51	68	173	2.08%
15	UNC Chapel Hill	49	55	61	165	1.99%
16	University of Texas at Austin	71	46	45	162	1.95%
17	Northwestern University	69	41	49	159	1.91%
18	UC Berkeley	52	49	53	154	1.85%
19	Cornell University	63	39	48	150	1.81%
20	University of Maryland	40	50	49	139	1.67%
21	University of Southern California	38	66	33	137	1.65%
22	UIUC	46	48	37	131	1.58%
23	ECGI	37	27	56	120	1.44%
24	Boston College	34	52	31	117	1.41%
25	Yale University	45	25	43	113	1.36%
26	University of Washington	25	61	24	110	1.32%
27	Washington University in Saint Louis	37	27	42	106	1.28%
28	Indiana University	33	37	32	102	1.23%
29	University of Rochester	28	55	15	98	1.18%
30	HKUST	31	37	25	93	1.12%
31	Federal Reserve Board	39	35	17	91	1.10%
32	Swiss Finance Institute	22	41	28	91	1.10%
33	University of Utah	26	40	25	91	1.10%
34	Arizona State University	28	45	17	90	1.08%

Table 1. - continued

Rank	Institutions	JF	JFE	RFS	All Journals	% of Total
35	INSEAD	18	33	37	88	1.06%
36	LSE	26	14	48	88	1.06%
37	Princeton University	38	33	16	87	1.05%
38	University of Notre Dame	35	37	15	87	1.05%
39	Purdue University	23	42	19	84	1.01%
40	Tilburg University	31	33	18	82	0.99%
41	University of British Columbia	27	20	34	81	0.98%

42	University of Oxford	36	22	23	81	0.98%
43	Emory University	23	38	18	79	0.95%
44	University of Toronto	31	25	23	79	0.95%
45	University of Minnesota	27	26	24	77	0.93%
46	University of Florida	20	39	14	73	0.88%
47	Michigan State University	14	28	30	72	0.87%
48	University of Virginia	24	24	24	72	0.87%
49	Carnegie Mellon University	23	14	34	71	0.85%
50	Federal Reserve Bank of New York	19	24	28	71	0.85%
Total		2927	2916	2463	8306	100%

Table 2. Summary of the 50 most prolific institutions. We categorize the institutions by type, geographic region, and time of establishment.

Type	Universities/Colleges	Research Bureaus/Centers	Banks&Others
Number in Total	45	3	2
Region	North America	Europe	Asia
Number in Total	42	7	1
Establishment	Before 1995	2002	2006
Number in Total	48	1	1

Table 3. Power law exponents for institutions' top finance journal publications. Estimates are from OLS regressions with standard errors in parentheses.

	All Journals	JF	JFE	RFS
α	1.640	1.490	1.689	1.615
	(0.328)	(0.298)	(0.338)	(0.323)
R^2	0.986	0.988	0.982	0.961

Table 4. Power law exponents for institutions' top finance journal publications over a 5-year period. Note the stability in *α* regardless of increasing mean number of publications.

Year	Mean Number of Publications	*α*	R^2
2015	155	1.632 (0.326)[b]	0.985
2014	146		0.984
		1.616 (0.323)	
2013	135	1.610 (0.322)	0.980
2012	123	1.572 (0.314)	0.973
2011	113	1.541 (0.308)	0.965

Table 5. Tests of the power law models to empirical data. β is the Gabaix- Ibragimov statistic with standard errors in parentheses.

	All Journals	JF	JFE	RFS
$\widehat{\beta}$	0.102	0.109	0.182	0.224
	(0.269)	(0.222)	(0.285)	(0.261)

Table 6. Power law models by Hill's tail-index estimator. α^H is the Hill's estimator with standard errors is in parentheses. KS-Stat is the Kolmogorov-Smirnovb test statistic with p-values is in parentheses.

	All Journals	JF	JFE	RFS
Panel A: Hill's tail-index estimator				
α^H	1.489	1.208	1.516	1.313
			(0.208)	(0.180)
(s.e.)	(0.205)	(0.166)		
Panel B: Kolmogorov-Smirnov test				
KS-Stat	0.065	0.131	0.123	0.124
(p-val.)	(0.968)	(0.300)	(0.373)	(0.357)

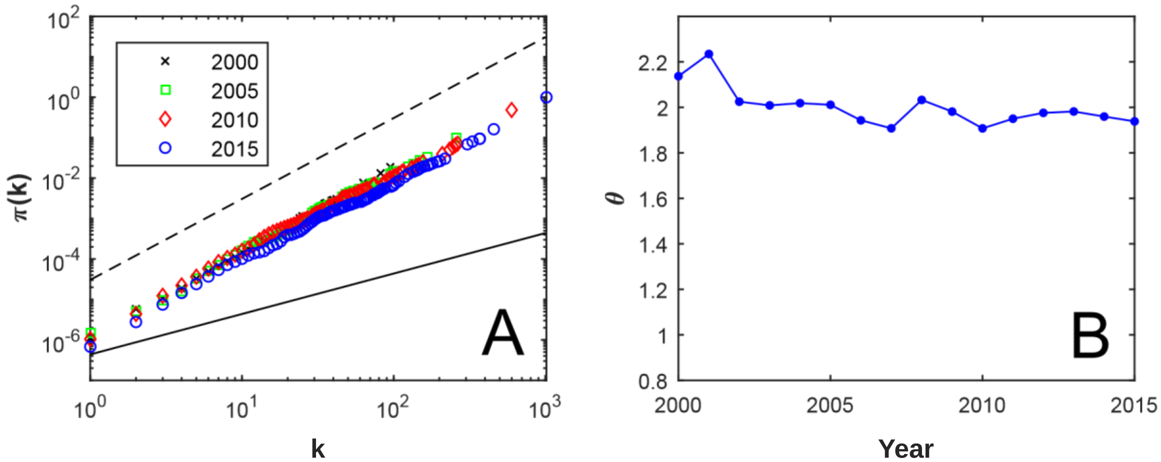


Figure A.1. The $\pi(k)$ function for the collaboration network (A) and the θ exponent (B). For each curve we used $\Delta t = 1$ *year*. We measure θ for each year from 2000-15 by fitting the whole $\pi(k)$ function. We plot the linear and quadratic $\pi(k)$ functions as the solid and dashed lines in log-log coordinates, respectively.