

# Informal collaboration in financial economics

## Final report

Patrick Altmeyer

26 March, 2021

Rose and Georg (2018) set out to analyse how informal collaboration among scholars affects research outcomes. To do so they use a dataset comprising meta data of more than 5000 articles from the academic literature on Financial Economics published between 1997 and 2011. Although they are hesitant to draw and causal connections, the authors find significant links between informal networking - conferences, seminars and paper reviews - and the success of papers and ultimately researchers. The estimated relationships are stronger than for formal co-authorship, which had been the primary focus of previous research in the line of literature. Perhaps not surprisingly then, more than half of the papers analysed in Rose and Georg (2018) involve all three kinds of informal collaboration. The findings set out in the paper are relevant not just to the hiring committees of academic institutions, but also young researchers at the beginning of their academic careers, who Rose and Georg (2018) find to more actively seek informal collaboration than their predecessors. Perhaps their methodological and data work is an even more valuable contribution. Their network of informal collaborations is publicly available and can be used to understand the drivers of scientific productivity. More importantly, their methodology is reproducible and in fact has already been applied to other fields of science (Dong et al. (2021) and Rose, Georg, and Opolot (2020)).

In this short note I will first briefly summarize the methodology and data used in Rose and Georg (2018). The following section then summarizes their main empirical findings (section ??). The final section offers a critical review of their work in context of the broader literature.

## Rose and Georg (2018)

### Data and methodology

The network of informal collaboration in Financial Networks is a novel data set derived from the acknowledgement sections of over 5000 published research papers from three top journals – The Journal of Finance, The Review of Financial Studies, the Journal of Financial Economics – and three second-tier journals – the Journal of Financial Intermediation, the Journal of Money, Credit & Banking, and the Journal of Banking and Finance. The raw data has been made freely available along with a short demonstration of how to link authors and commenters in a graph.

The networks they build go beyond the simple...

### Leading questions

Before turning to what we commonly understand as network analysis, Rose and Georg (2018) demonstrate the usefulness of their data through a number of simple regression models. The purpose of these empirical exercises is to shed light on a few leading questions: whether informal collaboration actually improves research, what type individuals engage (or are engaged) in the informal network and why researchers choose to form informal ties.

To be able to answer these questions Rose and Georg (2018) use the data underlying the network to derive a number of researcher characteristics including an individual's *prolificness*, experience and gender. The former is measured through a simple proxy measure based on their citations and publications. Experience is defined as the number of years between the first publication and the year of publication of the paper. Gender is estimated from authors names. Rose and Georg (2018) limit their attention to academic authors, omitting from the analysis any research assistants, non-academic advisers and editorial board members. The latter are removed from the analysis since their names can be expected to show up regularly in the acknowledgements section, simply due to their role as editors. Aside from researcher characteristics, the data comes equipped with a set of variables relating to the intensity of informal collaboration associated with individual papers. Those include the number of seminars attended, the number of conferences attended as well as the number of commenters.

With that data at hand the authors proceed to answer the leading questions empirically. With respect to the first question they estimate several modifications of the following model

$$\begin{aligned} \text{Success}_p = & \alpha_1 \text{Paper Characteristics}_{p,t-1} + \beta_1 \text{No. of seminars}_p + \beta_2 \text{No. of conferences}_p \\ & + \beta_3 \text{No. of commenters}_p + \beta_3 \text{Commenter Quality}_p + \mathbf{D}_{\text{Journal}_p} + \mathbf{D}_t + \varepsilon_p, \end{aligned} \quad (9)$$

where success of a paper  $p$  is measured in terms of its citation and acceptance by a top journal. Their estimates provide statistically significant evidence that with the exception of *conferences* all forms of informal collaboration contribute to the success of a paper. Descriptive statistics seem to demonstrate that it is mainly younger authors who increasingly rely on informal networking.

Turning to the second question about who engages in informal collaboration Rose and Georg (2018) estimate variants of the following model

$$\begin{aligned} \text{Collaboration}_{i,t} = & \beta_0 + \beta_1 \text{female}_i + \beta_2 \text{Euclid}_{it} + \beta_3 \text{PublicationStock}_{it} + \\ & \beta_3 \text{CitationStock}_{it} + \beta_4 \text{Experience}_{it} + \beta_5 \mathbf{D}_{\text{Female}_i} + \varepsilon_{it} \end{aligned} \quad (10)$$

where the dependent variable is a measure of how many times a reasearcher was included in the acknowledgements section of a paper. Only a minority of researchers acts as both authors and commenters at the same time, so this separation into producers and helpers of research as the authors describe it seems reasonable. The empirical estimates suggest that researcher charactersitics have significant impact on likelihood of being acknowledged. Most strikingly perhaps, gender appears to play an important role with female researchers up to 20% less likely to find their names in the acknowledgement sections.

Finally, with respect to the third leading question about researcher's incentives to seek informal ties, the authors simply look at distributional properties of the underlying sample. They find that there is a tendency to help researchers in return for past or future help: more than two thirds of all papers in the sample involve commenters commenting on the work of co-authors, their own commenters or department colleagues. Another possible driver of informal collaboration put forward by the authors revolves around inter-generational transfers: older, more experienced researchers may choose to comment in order to support younger colleagues. The sample period is too small to establish conclusive evidence, but sample distributions nonetheless reveal that commenters tend to have between 7 and 20 years of experience.

## Network analysis

In the context of network science, the part of Rose and Georg (2018) we turn to next is really the core of the paper. The authors first briefly present summary statistics for two different networks: firstly, the network of informal collaboration already introduced above and secondly a network of co-authorship. They then move on to perform a forecasting exercise in the spirit of Ductor et al. (2014) in order to gauge how well information derived from the two networks can forecast future productivity of researchers. Next, Rose and Georg (2018) also analyse the impact of network variables on the success of individual papers. Finally, they also investigate what drives centrality of researchers. We shall now briefly review each of these exercises.

**Two networks** Network of informal collaboration is much more dense than network of co-authorship. But density of both networks has decreased over time, which is attributed to a vast influx of researchers.

- In the network of informal collaboration, two researchers are connected with a weighted directed link whenever one acknowledges the other on a published paper in our dataset.

**Forecasting researcher’s productivity** Use different centrality measures: degree, eigen and betweenness. Focus is limited to “giant” component since centrality measures are not comparable across components (**DOUBLE CHECK. POTENTIAL SHORTFALL?**).

- the effort brought forward in equilibrium corresponds to someone’s eigenvector centrality.
- Since eigenvector centrality focuses on connectivity and influence only, but remains silent about the importance of a researcher for knowledge flows, we also study betweenness centrality

Correlation between eigenvector centrality and betweenness centrality is not high: it appears that researchers with high influence are not necessarily the ones that connects clusters of researchers.

- network of informal collaboration contains information not embedded in corresponding co-author networks. Put differently, co-author collaboration networks are not merely a proxy for informal collaboration. Informal collaboration networks also capture a larger share of the variance.

Argue that this should help identify promising candidates during the hiring process (based on the commenters they are connected to). **REVERSE CAUSALITY**

**Success of papers** Network centrality has impact on success of paper:

- first, the commenters’ network centrality contains information for both citation count and journal publications
- Second, and most importantly, models with centralities computed in the network of informal collaboration outperform those models with centralities computed in the co-author network.
- Finally, the centrality of authors and commenters matter for publication in crucial ways. Specifically, betweenness centrality and eigenvector centrality of authors and commenters in both networks are correlated in different ways with the dependent variables.
- This finding suggests that authors and commenters that connect different communities (high betweenness centrality) are less likely to publish in one of the top journals although their papers are cited more than the average of the journal

## Centrality of researchers

- More senior researchers are less influential in the network (lower eigenvector centrality), while more prolific researchers are more likely to connect otherwise disparate research communities (higher betweenness centrality).

*what is/are the main takeaway(s) from the paper?*

- researcher's position in this network is predictive of her future productivity and the scholarly impact of the papers she comments on
- We find positive but rather weak correlations (around 0.5) between scientific productivity measures and the number of acknowledged contributions.
- Female researchers are acknowledged considerably less
- find centralities in the network of informal collaboration to have higher explanatory power than centralities in the co-author networks.
- Conditional having an acknowledgement, the average top journal publication acknowledges almost twice as many commenters as the average non-top journal publication (10.8 vs. 6.2).

Co-authorship and conference participation has increased over time, while number of commenters and seminars has remained stable. It also appears that younger researchers have been more likely to acknowledge commenters and more generally make use of informal collaboration.

- Notably Ductor et al. (2014) show that an economist's future productivity can be forecasted using variables derived from a co-author network.

## Critical review

### Threats to validity

- Threats to validity: While strategic acknowledging is discouraged and can be costly, it remains the main threat to reliability measuring informal collaboration with acknowledgements data.

The rationale for excluding industry professionals and central bankers from the analysis is not very clear. After all, these professionals play an important role in academic research around Financial Economics: they attend and organise conferences, can often provide a unique first-hand perspective on financial markets and regularly produce research published in academic journals themselves.

### Related literature

While they do mention... - Too strong authority (or leadership) may have a negative effect, as it prevents theories to be challenged (Azoulay et al., 2019). Too much close networks might result in biased academic reviews (Carrell et al., 2020). ... the focus in their conclusion is

1. Rose, Georg, and Opolot (2020): Even among a set of papers that is highly selected on expected quality, discussants matter for top journal publication. Conditional on having a discussant, a paper's citation count increases in the discussant's prolificness. Using social network analysis we rule out a diffusion channel through which citations accumulate because discussants diffuse information about the paper within their social network.
2. Dong et al. (2021): show that preferential allocation of publications through success-breeds-success engenders disparity in institutions' research output, while acceleration in the growth of collaboration reduces the dispersion.

*is there anything you particularly like or dislike from the methodology/analysis? Would you suggest a different/complementary approach?*

*do you think there are open issues/questions that could be considered?*

*is there anything that could extrapolate to other situations as well?*

*have you found in the paper what you were expecting? If not, how does it differ from your expectations?*

## References

- Dong, Hui, Dan Luo, Xudong Zeng, and Zhentao Zou. 2021. “Cross-Affiliation Collaboration and Power Laws for Research Output of Institutions: Evidence from Top Journals in Financial Economics.”
- Ductor, Lorenzo, Marcel Fafchamps, Sanjeev Goyal, and Marco J Van der Leij. 2014. “Social Networks and Research Output.” *Review of Economics and Statistics* 96 (5): 936–48.
- Rose, Michael E, and Co-Pierre Georg. 2018. “What 5,000 Acknowledgements Tell Us about Informal Collaboration in Financial Economics.” *Max Planck Institute for Innovation & Competition Discussion Paper*, no. 11.
- Rose, Michael E, Co-Pierre Georg, and Daniel C Opolot. 2020. “Discussants.” *Max Planck Institute for Innovation & Competition Research Paper*, no. 20-19.