

Rose, Michael E.

**Article — Published Version**

## Small world: Narrow, wide, and long replication of Goyal, van der Leij and Moraga-González (JPE 2006) and a comparison of EconLit and Scopus

Journal of Applied Econometrics

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Rose, Michael E. (2021) : Small world: Narrow, wide, and long replication of Goyal, van der Leij and Moraga-González (JPE 2006) and a comparison of EconLit and Scopus, Journal of Applied Econometrics, ISSN 1099-1255, Wiley, Hoboken, NJ, Vol. 37, Iss. 4, pp. 820-828, <https://doi.org/10.1002/jae.2886>

This Version is available at:

<https://hdl.handle.net/10419/265080>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

**REPLICATION**

# Small world: Narrow, wide, and long replication of Goyal, van der Leij and Moraga-González (JPE 2006) and a comparison of EconLit and Scopus

**Michael E. Rose** 

Max Planck Institute for Innovation and Competition, Munich, Deutschland, Germany

**Correspondence**

Michael E. Rose, Max Planck Institute for Innovation and Competition, Marstallplatz 1, Munich, Deutschland, 80539, Germany.  
Email: Michael.Rose@ip.mpg.de

**Summary**

I undertake a narrow, wide, and long replication of Goyal, van der Leij and Moraga-González (2006, <https://doi.org/10.1086/500990>). Using social network analysis, they show that the Economics profession gradually evolved into a small world. Small worlds (or small world networks) have unique information transmission capabilities. The trend is explained by the emergence of frequently publishing researchers with many distinct co-authors. In a social network, they resemble stars. The original results are robust to the usage of (I) another software, (II) a recent version of the originally used data, and (III) another database and a more sophisticated author disambiguation.

**KEYWORDS**

economics, social network, small world

## 1 | INTRODUCTION

Economists are driven by a strong desire to understand the world, where, tautologically, the Economics profession is part of that world. A rich literature stream thus studies the functioning, impact, and design of the Economics profession. Recent prominently published examples include Fourcade et al. (2015), Angrist et al. (2017), Heckman and Moktan (2020), and Hamermesh (2018). The literature reviews by Colander (1989) and Coupé (2004) testify of this long-standing interest. Understanding knowledge flows and spillovers among collaborators, for instance, is an important topic in this regard and can be addressed by social network analysis. Social network studies have thus been at the early forefront of this literature, and Goyal et al. (2006) is the pioneering study of social network analysis in the Economics of Economics.

This paper both narrowly and widely replicates Goyal et al. (2006). The authors analyze the economics profession in the 1970–1999 period and find that by the 1990s, Economics has become a small world. I succeed to replicate their findings in gradually increasing degrees of alterations to the setting of the original study: using a different software, using a recent excerpt from the original database, and using an alternative database along with a different author disambiguation. In all instances I fully replicate the findings of Goyal et al. (2006). Additionally, I undertake a long replication using data on two more decades.

The original data stem from EconLit, a specialized database for Economics literature only and curated by the American Economic Association.<sup>1</sup> My alternative database is Scopus, which is generally more extensive, has a different coverage (of

<sup>1</sup>See <https://www.aeaweb.org/econlit/>

the journals covered in EconLit), and has a more sophisticated author name disambiguation procedure (Baas et al., 2020).

By comparing two established bibliometrics databases on a single application, especially in the long replication, this study has a further contribution. It helps the scientometrics community to understand the impact of the use of different bibliometrics databases. These differences matter because, after all, they are used in science assessments and potentially affect scientists' careers.

## 2 | ECONLIT VERSUS SCOPUS

Scopus is a widely used citation and abstracting database provided by the publisher Elsevier. It covers all sciences and provides the underlying data in many assessments, such as The Times Higher Education rankings, and gaining in popularity among scientists. For this reason I have picked Scopus as database for the wide replication.<sup>2</sup> Scopus has furthermore three advantages over EconLit, but also a disadvantage, which I discuss in this section.

While Scopus is very broad, the originally used EconLit covers the one narrow discipline that it covers to a much deeper extend. Between 1970 and 1999, the period of study of Goyal et al. (2006), EconLit indexes 850 different journals (accounting for name changes) and additionally 13 (for the wide replication) resp. 273 (for the long replication) articles published in the Proceedings of the National Academy of Sciences (PNAS).<sup>3</sup> Scopus indexes 998 out of the 1748 EconLit journals for at least 1 year during the study period (as well as all PNAS articles). Thus, EconLit beats Scopus in terms of coverage of journals.

Coverage of a journal however does not imply perfect coverage from the first year of operation until today. Indeed, coverage may start later or end earlier or—in the case of Scopus—may have considerable gaps.<sup>4</sup> In total, EconLit (as of today) covers 10,190 volumes in the 1970–1999 period (I used the actual coverage in the data, not the one indicated online). Figure 1a shows that there are many volumes covered in EconLit but not in Scopus (4901) but also few volumes covered in Scopus but not in EconLit (2326). In the wide replication I use the 5289 volumes indexed in both databases, as well as the 13 PNAS articles. This effectively reduces the analysis to 52% of the volumes of EconLit. In the 2000–2019 period, EconLit covers 20,546 volumes (Figure 1b). In the long replication with Scopus I use 12,278 (59.8%) volumes that are indexed in Scopus and in EconLit.

Yet Scopus offers three features that EconLit does not. The first and most important for network studies are entities (Baas et al., 2020; Moed et al., 2013). Scopus strives to maintain exactly one Scopus author profile for every author, and each profile has a unique ID on which the researcher can rely. The assignment is based on name, field, and affiliation.<sup>5</sup> While Scopus does not disclose its disambiguation algorithm, it is reasonable to assume that is more sophisticated than what individual scientists come up with on their own. What is known is that the algorithm prefers “split profiles” over “merge profiles” (Moed et al., 2013): It rather casts too many profiles for the same researcher than accidentally lumping publications of different researchers into one profile. Accordingly, Baas et al. (2020) estimate precision of Scopus author profiles equal to 98.1% and recall equal to 94.4%. Goyal et al. (2006) disambiguate based on surname and first initial(s) only. As the authors now, this algorithm favors recall over precision. However, even recall might not be high since a different order of first names (e.g., Yun Jack vs. Jack Yun) results in two different authors. To estimate downside of this, I investigate the right tail of the publication count distribution. This is where misallocations become visible the clearest. Figure 2 compares the number of publications in the upper 1% of the publication count distribution. The left (brown) boxplots represent the EconLit networks; they are always longer than the purple Scopus networks. The difference aggravates over time, but it does not look suspicious in the first three decades. The highest publication count in the 1990s EconLit network equals 201 and belongs to “creedy j.” With the emergence of Asian scholars though in the 2000s, new short and frequent names make it to the publication sphere, and the algorithm creates authors with enormous publication counts: In the last two decades, the top three authors according to EconLit in the 2000s are called “lee j” (201 publications), “zhang j” (169),

<sup>2</sup>Although, the term “wide” might not apply perfectly: Both databases cover in principle the same publications. It's the coverage, the processing and the aggregation that makes them different.

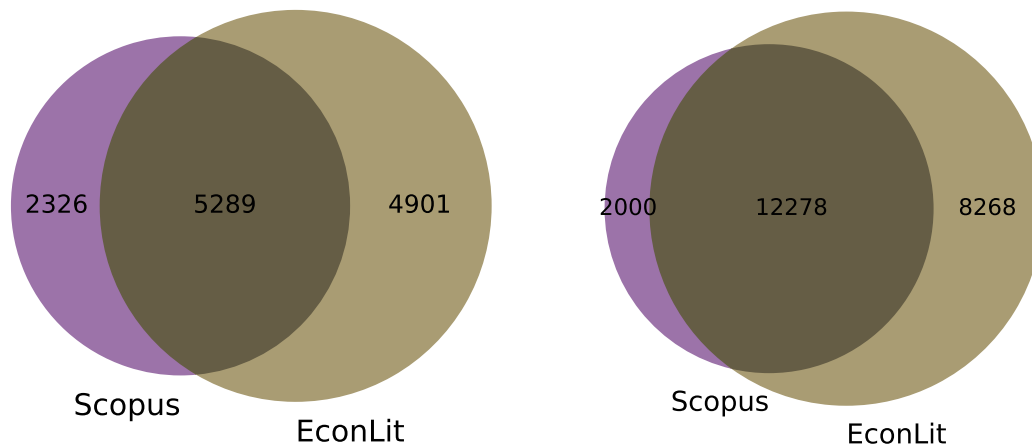
<sup>3</sup>The EconLit I use stems from December 2020, which differs from the originally used EconLit. It indexes few more volumes of indexed journals, excludes some low-quality journals altogether, and includes some new articles of already indexed volumes.

<sup>4</sup>These are alternating sequences of years of coverage of a journal although the journal operates. For example, as of December 2020, Scopus covers the Journal of Political Economy in 1969, from 1973 to 1974, from 1979 to 1985, and finally from 1987 to Present. Contrary to popular belief these coverage gaps are unintentional. Often the databases simply did not receive all meta data, which had to exist in digital form prior to the transfer.

<sup>5</sup>Corrections requested by users of Scopus (typically merging of some split profiles) can enhance the profiles' accuracy further.

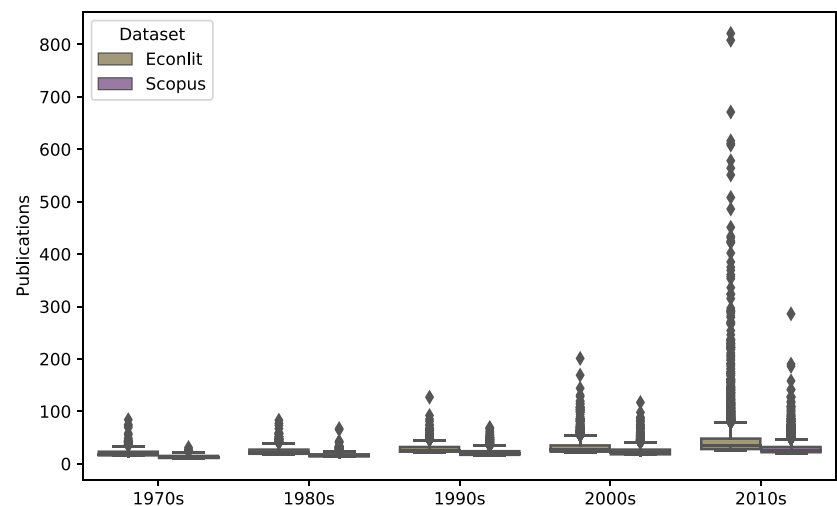
(a) 1970s, 1980s and 1990s.

(b) 2000s and 2010s.



**FIGURE 1** Volumes of EconLit journals covered in EconLit and Scopus. *Note:* The brown circles show the number of *volumes* (=journal-year observations) covered in EconLit of all journals actually indexed in EconLit as of 2020. The purple circles show the number of *volumes* actually indexed in Scopus as of 2020. For the narrow replication, we use the brown circle and for the wide and long replications, we use the intersection

**FIGURE 2** Comparison of upper tail of publication count by network. *Note:* Boxplots visualize the distribution of the upper 1% of the publication count distribution by network



and “wang y” (144) and in the 2010s their names are “zhang j” (821), “zhang y” (808), and “wang y” (671). By contrast, the highest publication count in the 2000s Scopus network equals 286.<sup>6</sup>

We can paint the same picture studying the distribution of the count of unique co-authors (degree) instead. In the 1990s, the highest degree is 54 according to EconLit and 65 according to Scopus. But in 2010 already, the highest degree equals 968 and belongs to “zhang y.” There are 779 individuals with degree larger than 54, almost all of which have Asian names. By comparison, the highest degree in the 2010s Scopus network equals 193, and there are only 402 authors with degree larger than 54. Overall, one can expect the EconLit networks to overstate the importance of well-connected authors and to bias towards common surnames in the 2000s and in the 2010s, but for the first three decades there is probably no such bias.

The second advantage of Scopus over EconLit is how one accesses the data. Scopus maintains a public RESTful Automated Programmer Interface (API) for subscribers (Baas et al., 2020). The API ensures data integrity as the data are pulled from the origin directly and not sent around in text files; it eases data retrieval, and it facilitates replicability (Rose & Kitchin, 2019).

<sup>6</sup>This profile belongs to Rangan Gupta from University of Pretoria with 426 documents as of end of 2020. A comparison with the publicly available CV indicates that the Scopus profile is correct.

Finally, EconLit reports only the first author of papers with four or more authors. In the original study, 1.6% of all the publications were by four or more authors, which Goyal et al. (2006) have dropped. In the recent EconLit, articles with four or more authors are not distinguishable from solo-authored papers, which is why I cannot drop them. Using Scopus however, I am able to include all these papers and the resulting connections.

### 3 | RESULTS

In this section I replicate Goyal et al. (2006) thrice. In the first subsection I use both the original dataset used by the authors as well as my own recent extract of the EconLit database. In Section 2, I use Scopus for the wide replication. Finally I present a long replication with recent EconLit and Scopus. I use exclusively Python 3.7, whereas Goyal et al. (2006) have used C++. For the network analysis I use NetworkX version 2.4 (Hagberg et al., 2004).

#### 3.1 | Narrow replication

To narrowly replicate the results of Goyal et al. (2006), I use the exact same publications data (EconLit publication lists for 1970–1999) and assign them to three groups corresponding to the three decades. The first main result of Goyal et al. (2006) is that only by the 1990s, the network of Economists satisfies all four criteria for the small world hypothesis and that academic Economics is an emerging small world. These criteria are as follows: Average number of co-authors is small relative to the total number of economist, there exists a sizeable largest component (giant component), distance within the giant component is small, and clustering is high (as compared with a random network with same degree). The network statistics, shown in Table 1, match exactly those in Table 1 of Goyal et al. (2006), with the exception of an additional isolated author in the 1990s network. Thus, the results are robust to using Python instead of C++.

Since there are multiple steps in the data processing, the more interesting case is of course to replicate the data gathering as well. To this end, I use the EconLit data as of June 2020. I use only research articles published in academic journals. I cannot however drop publications with four or more authors because only the first author is listed for such articles, and papers are indistinguishable from solo-authored papers. In Table 2, I again replicate the first main result of Goyal et al. (2006). The numbers are not always identical yet very very close. For example, I find 80 more authors in the 1970s, 258 more in the 1980s, and 463 more in the 1990s. The other network statistics differ mostly only in the digits.

The second main result of Goyal et al. (2006) is that highly interlinked Economists are the reason for the low (and decreasing) average distance and the high (and increasing) connectedness. By connecting other distant groups, highly interlinked Economists—so called “stars”—drastically reduce the distance between members of these groups. This translates into a low clustering coefficient for them. In Table 3 and Figure 3, I replicate Table 2 and Figures 1 and 2 of Goyal et al. (2006). Table 3 gives the number of papers, the share of them that are co-authored, the degree, degree of order 2 (number of unique co-authors of all immediate co-authors), and the clustering coefficient for the five most connected

	1970s	1980s	1990s
Total authors	33,770	48,608	81,216
Degree			
Average	.894	1.244	1.672
Standard deviation	1.358	1.765	2.303
Giant component			
Size	5,253	13,808	33,027
Percentage	15.6%	28.4%	40.7%
Second largest component	122	30	30
Isolated authors			
Number	16,735	19,315	24,577
Percentage	49.6%	39.7%	30.3%
Clustering coefficient	.094	.142	.200
Distance in giant component			
Average	12.86	11.07	9.47
Standard deviation	4.03	3.03	2.23

TABLE 1 Network statistics for the co-author networks using original EconLit

Note: Narrow replication of Table 1 of Goyal et al. (2006) using their data.

**TABLE 2** Network statistics for the co-author networks using recent EconLit

	1970s	1980s	1990s
Total authors	33,850	48,866	81,680
Degree			
Average	.889	1.237	1.652
Standard deviation	1.349	1.759	2.298
Giant component			
Size	5,146	13,768	32,879
Percentage	15.2%	28.2%	40.3%
Second largest component	122	30	30
Isolated authors			
Number	16,816	19,512	25,380
Percentage	49.7%	39.9%	31.1%
Clustering coefficient	.093	.141	.197
Distance in giant component			
Average	12.99	11.13	9.48
Standard deviation	4.16	3.05	2.22

Note: Narrow replication of Table 1 of Goyal et al. (2006) using EconLit obtained in June 2020.

**TABLE 3** Network statistics for most linked economists using recent EconLit: 1990s

	Papers	% Co-authored	Degree	Distance 2	Clustering coefficient
1	66	97.0	54	243	.022
2	58	56.9	43	154	.020
3	67	100.0	41	172	.045
4	68	92.6	41	57	.034
5	48	93.8	34	168	.036
Average top 100	38.41	81.0	25.30	96.63	.042
Average all	2.82	59.9	1.65	5.17	.157

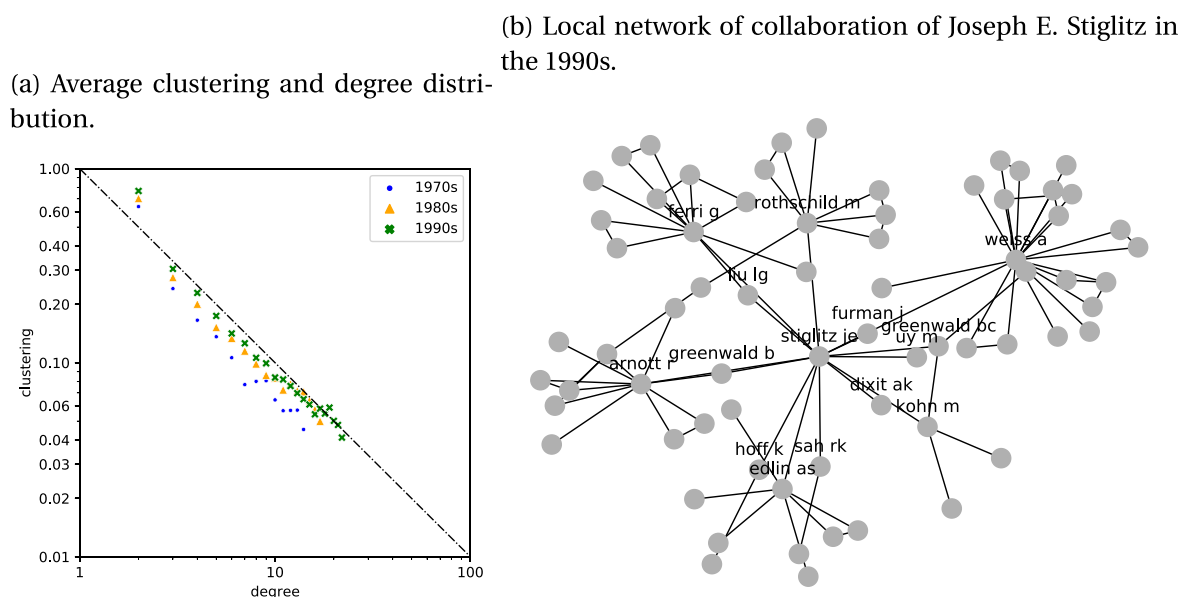
Note: Narrow replication of Table 2 of Goyal et al. (2006) using EconLit obtained in June 2020. The network is composed of “star” Economists that have many co-authors but a low clustering coefficient.

Economists. Half of these figures match precisely the original table; the other half is off by just few percent. The story behind these numbers is that highly productive authors are those with low clustering coefficient, that is, they write many papers with many co-authors that do not co-author with each other. Figure 3a corroborates this relationship graphically. Except for some high-degree individuals, my dots mirror those of Goyal et al. (2006). Finally, Figure 3b exemplifies this relationship with the co-author network of Joseph E. Stiglitz. There are some differences between Figure 3b and its counterpart, which are likely due to some undocumented manual corrections of names, such that “dixit a” becomes “dixit ak.”

### 3.2 | Wide replication

I widely replicate the results using the Scopus database as far as it covers the volumes indexed in EconLit. I pull all the data using code developed by Rose and Kitchin (2019) in December 2020. There are two important changes on how I define the network. First, I include all pairwise co-author connections from papers with four or more authors. Second, I do not disambiguate author names myself to construct network nodes but rely on Scopus' definition of an author. That is, I use Scopus author profile IDs. This disambiguation can deal with the overreporting (same author with different names) and underreporting (two authors with the same name lumped together) inherent to the disambiguation procedure of Goyal et al. (2006).

As expected, the networks differ. Table 4 reports considerably fewer authors in the networks, but the gap narrows over time. Yet except for the 1970s, the share of the giant components is very similar. Average clustering coefficient is up by about 50% as compared to Table 3. Despite these differences, by the 1990s, the network satisfies all criteria of a small world.



**FIGURE 3** Clustering and degree using recent EconLit. *Note:* Narrow replication of Figures 1 and 2 of Goyal et al. (2006). Clustering for a given degree  $k < \bar{k}$  ( $\bar{k}$  is 14 in the 1970s, 17 in 1980s, and 22 in the 1990s) is the clustering coefficient of nodes with degree  $k$ . Observations for degree  $k \geq \bar{k}$  are grouped together

	1970s	1980s	1990s
Total authors	19,215.000	32,317.000	69,195.000
Degree			
Average	.983	1.777	2.258
Standard deviation	1.392	5.649	3.386
Giant component			
Size	1,076.000	8,544.000	28,762.000
Percentage	5.6%	26.4%	41.6%
Second largest component	80	36	43
Isolated authors			
Number	8,698	10,514	16,632
Percentage	45.3%	32.5%	24.0%
Clustering coefficient	.143	.215	.309
Distance in giant component			
Average	19.09	12.81	10.55
Standard deviation	8.75	3.87	3.08

**TABLE 4** Network statistics for the co-author networks using Scopus

*Note:* Wide replication of Table 1 of Goyal et al. (2006).

The five Economists with highest degree have fewer publications and higher clustering (Table 5) than in Goyal et al. (2006).<sup>7</sup> Also, in this network, I find a strong negative correlation between the number of co-authors and clustering of a researcher (Figure 4a), but it is not as negative as in the EconLit networks.

### 3.3 | Long replication

With the benefit of nearly two decades that passed since the publication of Goyal et al. (2006), we can undertake a long replication. I repeat the exact same analysis as above using all publications from the 2000s and 2010s.

<sup>7</sup>An investigation of the 5 most connected Economists reveals that Scopus and EconLit produce different rankings. According to EconLit (with above author disambiguation), the 5 most connected Economists are “thisse jf”, “lee j”, “sirmans cf”, “nijkamp p”, and “michel p”. According to Scopus these are “Costanza, Robert”, “Perrings, Charles A.”, “Wilen, James E.”, “Kellert, Stephen R.”, and “Smith, V. Kerry”. These are Ecologists whose publications are largely not covered in Goyal et al. (2006) because of the author group size restriction.

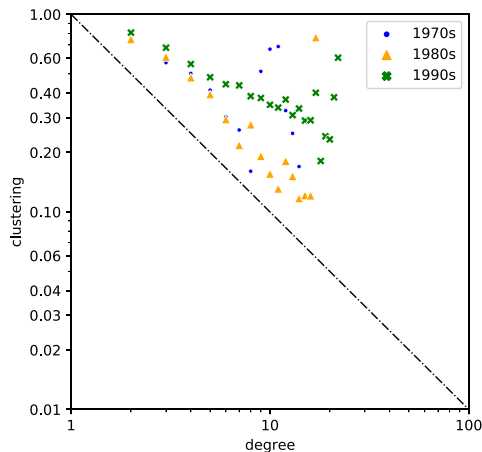


**TABLE 5** Network statistics for most linked economists using Scopus: 1990s

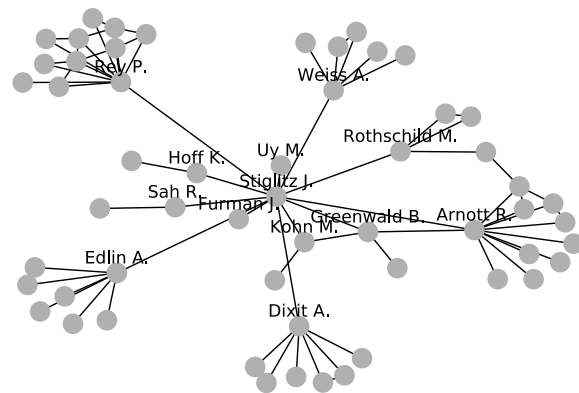
	Papers	% Co-authored	Degree	Distance 2	Clustering coefficient
1	15	80.0	65	164	.167
2	16	50.0	65	151	.468
3	12	91.7	61	170	.466
4	2	100.0	55	121	.614
5	37	70.3	52	241	.097
Average top 100	4.39	90.4	43.18	54.05	.884
Average all	2.39	69.1	2.25	6.30	.522

Note: Wide replication of Table 2 of Goyal et al. (2006).

(a) Average clustering and degree distribution.



(b) Local network of collaboration of Joseph E. Stiglitz in the 1990s.



**FIGURE 4** Clustering and degree using Scopus. Note: Wide replication of Figures 1 and 2 of Goyal et al. (2006)

**TABLE 6** Long replication of network statistics using EconLit

	2000s	2010s
Total author	157,712	301,591
Degree		
Average	2.524	3.976
Standard deviation	3.807	9.812
Giant component		
Size	84,852	203,471
Percentage	53.8%	67.5%
Second largest component	73	53
Isolated authors		
Number	33,301	36,806
Percentage	21.1%	12.2%
Clustering coefficient	.322	.446
Distance in giant component		
Average	7.91	6.64
Standard deviation	1.94	1.70

The networks of Economists in the 2000s and 2010s are much larger, owing to the increased amount of publications. Table 6 shows that the networks derived from EconLit double in size with each decade. Importantly, the giant components and the clustering components grow, while average distance in the giant component decreases. The world of Economists, which has been small since the 1990s, became even smaller.



	2000s	2010s
Total authors	132,338	286,295
Degree		
Average	2.991	3.939
Standard deviation	5.340	5.542
Giant component		
Size	66,724	182,471
Percentage	50.4%	63.7%
Second largest component	52	81
Isolated authors		
Number	22,913	27,723
Percentage	17.3%	9.7%
Clustering coefficient	.393	.513
Distance in giant component		
Average	9.15	8.55
Standard deviation	2.25	1.97

TABLE 7 Long replication of network statistics using Scopus

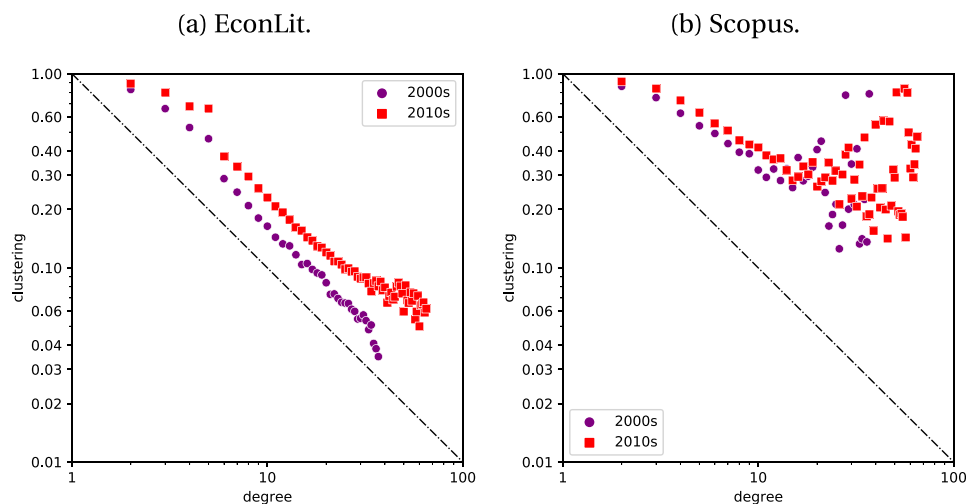


FIGURE 5 Long replication of the relationship between clustering and degree. Note: Long replication of Figure 1 of Goyal et al. (2006) using EconLit data (left) and Scopus data (right). Clustering for a given degree  $k < \bar{k}$  ( $\bar{k}$  is 37 in the 2000s and 65 in the 2010s) is the clustering coefficient of nodes with degree  $k$ . Observations for degree  $k \geq \bar{k}$  are grouped together

Table 7, which uses the Scopus networks, paints a very similar picture. The only difference is that the small world of Economists is somewhat less small: Average distance is higher and the giant component is a little less inclusive.

A remaining question is that of the clustering–degree relationship in the 2000s and 2010s. Figure 5 tackles this question. With EconLit (Figure 5a), the relationship of clustering to degree is not following closely an exponential relationship anymore. In this figure, high-degree individuals are grouped together if their degree exceeds a certain threshold  $\bar{k}$ . Goyal et al. (2006) chose three values for their three decades (14, 17, and 22). In the absence of a justification for  $\bar{k}$ , I define  $\bar{k}$  for the 2000s and 2010s such that for no  $k < \bar{k}$  there are fewer than 10 observations. This condition holds for the original values for  $\bar{k}$  as well.

Using Scopus data (Figure 5b), we see the same pattern as with EconLit data up until a degree of 20, after which however the pattern becomes nontrivial. There are high-degree individuals with high clustering, and there are high-degree individuals with low clustering. Clearly, a lower  $\bar{k}$  results in a clearer picture for high-degree individuals.

Nevertheless, EconLit and Scopus differ distinctly in that part of the distribution. One reason for this difference is the name disambiguation. The procedure for EconLit authors yields few profiles with an unusual high number of publications, which correlates with a high degree. As these co-authors are not authors of the same common researcher in reality, they are also not co-authors of each other in the network (at least not with the same probability as co-authors of low-degree authors). Thus, we are very likely to underestimate the true clustering of high-degree authors. Another reason are papers with many authors, which create author groups with perfect clustering. They are common in Ecological Economics and present in Scopus but exempt from EconLit (due to the author group size restriction).

## 4 | CONCLUSION

I replicate all main results of Goyal et al. (2006) thrice: using their data and their network definition, using a recent version of their data and their network definition, and using another dataset and a different network definition. A comparison between EconLit and Scopus shows that EconLit has the larger coverage (in Economics and neighboring fields), but Scopus trumps with its sophisticated author disambiguation.

## ACKNOWLEDGMENTS

For extraordinary research assistance, I wish to thank Nurzhan Sapargali. For helping clarifying issues, I thank Lorenzo Ductor and Marco van der Leij. I am furthermore grateful to two anonymous referees and the managing editor, Heather Anderson.

## ORCID

Michael E. Rose  <https://orcid.org/0000-0002-4128-4236>

## REFERENCES

- Angrist, J., Azoulay, P., Ellison, G., Hill, R., & Lu, S. F. (2017). Economic research evolves: Fields and styles. *American Economic Review: Papers & Proceedings*, 107(5), 293–297. <https://doi.org/10.1257/aer.p20171117>
- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377–386. [https://doi.org/10.1162/qss\\_a\\_00019](https://doi.org/10.1162/qss_a_00019)
- Colander, D. (1989). Research on the economics profession. *Journal of Economic Perspectives*, 3(4), 137–148. <https://doi.org/10.1257/jep.3.4.137>
- Coupé, T. (2004). What do we know about ourselves? On the economics of economics. *Kyklos*, 57(2), 197–215.
- Fourcade, M., Ollion, E., & Algan, Y. (2015). The superiority of economists. *Journal of Economic Perspectives*, 29(1), 89–114.
- Goyal, S., van der Leij, M. J., & Moraga-González, J. L. (2006). Economics: An emerging small world. *Journal of Political Economy*, 114(2), 403–412. <https://doi.org/10.1086/500990>
- Hagberg, A., Schult, D., Swart, P., & Networkx developers (2004). NetworkX. <https://networkx.github.io/documentation/stable/index.html>
- Hamermesh, D. S. (2018). Citations in economics: Measurement, uses, and impacts. *Journal of Economic Literature*, 56(1), 115–156. <https://doi.org/10.1257/jel.20161326>
- Heckman, J. J., & Moktan, S. (2020). Publishing and Promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, 58(2), 419–470. <https://doi.org/10.1257/jel.20191574>
- Moed, H. F., Aisati, M., & Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics*, 94(3), 929–942. <https://doi.org/10.1007/s11192-012-0783-9>
- Rose, M. E., & Kitchin, J. R. (2019). Pybliometrics: Scriptable bibliometrics using a Python interface to Scopus. *SoftwareX*, 10(July-December), 100263.

**How to cite this article:** Rose, M. E. (2022). Small world: Narrow, wide and long replication of Goyal, van der Leij and Moraga-González (JPE 2006) and a Comparison of EconLit and Scopus. *Journal of Applied Econometrics*, 37(4), 820–828. <https://doi.org/10.1002/jae.2886>