International Journal of Accounting and Financial Reporting
ISSN 2162-3082
2024, Vol. 14, No. 2

The Most Prestigious Research Areas in Financial Economics From 1896 to 2006: Scientometrics Based on Complex Networks

Marcelo De Oliveira Passos

Federal University of Pelotas, Brazil

Jéssica Rodrigues Gonçalves

Federal University of Pelotas, Brazil

Mathas Schneid Tessmann (Corresponding author)

Brazilian Institute of Education, Development and Research – IDP, Brazil

E-mail: mathias.tessmann@idp.edu.br

Daniel De Abreu Pereira Uhr

Federal University of Pelotas, Brazil

Alex Cerqueira Pinto

University of Brasília, Brazil

Received: April 17, 2024 Accepted: May 20, 2024 Published: May 24, 2024

doi: 10.5296/ijafr.v14i2.21841 URL: https://doi.org/10.5296/ijafr.v14i2.21841

Abstract

This work has three objectives related to scientometrics of financial economics from 1896 to 2006: (i) to detect which are the most cited authors and co-authors of a sample of the most influential works in the finance literature; and (ii) define the most relevant co-authorship groups in this sample; and (iii) develop a complex network with the links between these clusters, authors and co-authors. We used the Complex Network Statistics weighted degree metric, IDEAS/RePEc scores, and ranking to achieve the first objective. For the second, we adopt the modularity class process. For the third, we use Yifan Hu's proportional layout algorithm. The database was gathered from two sources, the Institute for New Economic Thinking's The History of Economic Thinking website and the references described by financial historian Peter L. Bernstein in his seminal book tracing the history of financial and economic thought.

Keywords: Scientometrics, Financial economics, Network analysis, Modularity classes, Financial history

JEL Classification: C80, B26, E44, N20

1. Introduction

Topics related to the financial economy have increasingly attracted society's attention. Due to more agile and intense information flows, more people have begun to invest their resources in the capital market in recent years. It may be the main reason for the increase in this generalized interest. Likewise, the literature that investigates the subject has become proportionately more consulted because it is used as a reliable source by investors who seek a better understanding of the functioning of the markets. Since the 1950s, with the initial study published by Markowitz on the theory of portfolio optimization, the scientific literature that investigates financial economics has constantly been evolving, undergoing diverse and equally essential advances, such as Sharpe's development, in 1964, on asset pricing. Moreover, the efficient market hypotheses, the introduction of agency costs by Jensen and Meckling in 1976, and the principles of behavioral economics presented by Thaler, Tversky, and Kahnemann in the 70s and 80s.

Analyzing how all these advances took place through identifying the most prestigious themes in finance is important because it helps the interested community understand its evolution in a more complete and structured way. For this, the most efficient strategy is scientometrics, which is dedicated to investigating the quantitative aspects of how Science is produced, considering such production as a complex system (complex network, in the case of this work) of communication. One of its main focuses, but not the only one, is the analysis of citations in the academic literature.

Scientometrics has been expanding as an area of high relevance for measuring and evaluating several research areas with the evolution of data science and the popularization of data collection software in large databases. Scientometrics is directly linked to the investigation and evaluation of scientific research. However, it has some intrinsic limitations. It can quantify and analyze book loans in different libraries, but there is no way for it to assess the effective reading of books and periodicals in libraries, in loco. It can also measure online access and downloads, leading to webometrics and altmetrics, areas beyond the scope of this article. Quotations from other authors can be understood as links between the intellectual productions of people, institutions, or electronically produced data sources. Such links can be evaluated in the context of a complex network which, in turn, can be analyzed visually and statistically. The same occurs with works carried out in co-authorship. However, the bond between the co-authors is stronger than a simple citation of an author external to the work produced by them.

The objectives of this work are three: (i) to identify which were the most cited authors and co-authors of a sample of the most influential articles in the finance literature; (ii) define the most relevant clusters of co-authorship in this sample; and (iii) to elaborate a complex network in which it is possible to visualize what these clusters are and, at the same time, graphically identify which are the most relevant in literary production on financial economics.

When referring to scientific literature, it is important to note that Google Scholar has increased its scope over the years and is a great database for scientific literature (Moed and Bar-Ilan, 2017) at the same time as the Research Papers in Economics network. (RePEc) is the largest bibliometric database in economics (Seiler and Wohlrabe, 2014). Thus, several works have made important contributions to the financial economics literature, the findings of Seiler and Wohlrabe (2014) show that the impact factors of the literature in question are commonly reliable, Galiane and Gálvez (2019) identified that in the financial literature there is no there is a spike in annual citations for the same age as the most cited papers and Franceschet (2010) found that five finance and economics journals are among the most prestigious in the social sciences.

The work by Amon and Hornik (2022) used text mining to verify the recurrence of characteristics of the most prestigious papers and concluded that text length, international collaboration, personal and relatively informal writing style, and paper density are among the main predictors of prestige, in Wu et. Al (2018) identified that the positioning in the ranking of research institutions in economics is influenced by the prestige of the journals to which their affiliates make up the editorial team, and that of Tol (2013) found that researchers with a greater number of citations end up receiving more citations. disproportionate to the quality of their work. Thus, we believe that this work fills a gap in the scientometric and financial economics history literature. There is still very little research that uses statistical methods of complex networks capable of quantifying, classifying, and describing in detail the most relevant lines of research and authors in an area of scientific knowledge.

In addition to this introduction, the article consists of four more sections. Section two presents a brief review of the literature on graphs, networks, and scientometrics, section three describes the methods and data used, section four presents the results, and, finally, section five concludes.

2. Theoretical Reference

2.1 The Basics of Graph and Network Theory

According to Easley and Kleinberg (2012), a graph is a way of specifying relationships between a collection of items. A graph consists of a set of objects called nodes, some of which are connected by edges. Figure 1 below shows examples of two types of graphs, undirected graph (a) and directed graph (b), each with four nodes and four edges. In figure (1.a), "A" is connected with "B" and "B" is connected with "A." In Figure (1. b), "A" is connected with "B," but "B" is not connected with "A."

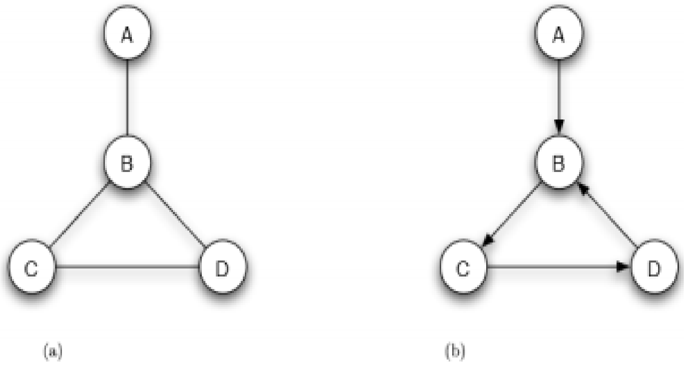

Figure 1. (a) Undirected graph (b) Directed graph (digraphs)

Source: Easley e Kleinberg (2012).

Graphs are helpful as they can be used as a mathematical network structure model. An example of the structure class of a graph is the social network, where the nodes are people or groups of people and the edges are a type of social interaction. Another example is an information network, where the nodes can be documents, and the edges can be references between documents.

There are empirically created graph models with specific characteristics that identify the types of graphs that can exist. In addition, the most diverse existing metrics can be used to calculate the characteristics of a graph. The attributes most commonly used in previous studies in other countries were selected to characterize a citation network.

2.2 Types of Networks

Empirically studied models help in the study of real networks; each type of network has statistical properties to identify. Among the types of networks that exist are random, small-world, and free-scale networks.

2.2.1 Random Networks

Mathematicians Erdös and Rényi (1959 and 1960) began to study networks and random graphs. In this model, they generated a network with N nodes in which they connected each pair of nodes with a probability p, thus generating a random distribution of edges of approximately $pN(N-1)/2$. This model has been used as a basis for research into complex networks for decades, as real networks were believed to lack organizational principles. Nevertheless, this argument was overcome with time.

Random networks are important because they can be used as a helpful comparison model for analyzing real complex networks when we know their intrinsic properties.

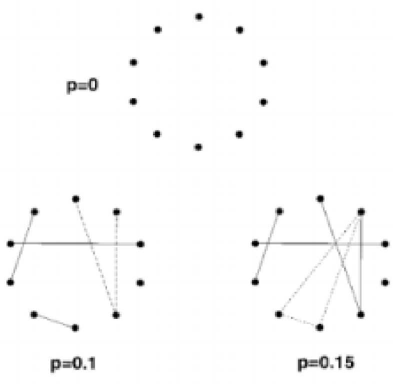

Figure 2. Random graphs with different probabilities of connections between nodes

Source: Albert and Barabási (2002)

2.2.2 Small-World Networks

Small-world networks are a graph in which most connections are established between the closest vertices. The main characteristic of small-world networks is that the path between any pair of nodes in the network is relatively short. The path is represented by the minimum number of edges from one node to another. Stanley Milgram developed the theory of six degrees of separation in 1967. Based on their social relationships, any United States resident would have six edges (or connections or contacts) far from any other North American (Barabási and Albert, 1999).

Barabasi and Albert (1999) and Barabási and Bonabeau (2003) argue that small-world networks cannot indicate an organizing principle, as random graphs can have this characteristic. Figure 3 below shows an example comparing conventional networks, small-world networks, and random networks. It is worth remembering that regular networks are those in which all nodes have the same degree.

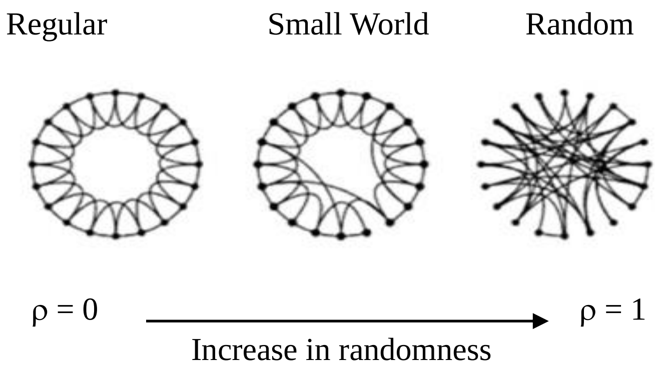

Figure 3. Comparison of edge distribution among regular type graphs

Source: Albert and Barabási (2002).

Note: The distribution of edges varies according to the probability p-value of a node connecting to another.

2.2.3 Scale-Free Networks

Albert and Barabási (2002) studied free-scale networks, which, despite following an exponential distribution, often also differ from random networks because they do not have a degree distribution that follows the Poisson distribution.

Free-scale networks are characterized by a few nodes having many connections and many nodes having few connections. In addition, a new node that appears in the network is more likely to connect to a node with more connections than one with fewer. Barabási and Bonabeau (2003) cite existing real examples, such as the World Wide Web, where sites with many citations tend to continue to be widely cited over time, as well as scientific collaborations.

2.2.4 Complex Networks

Complex networks have their origins in graph theory. However, graph theory was initially focused on the study of regular graphs, and large-scale/complex networks, since 1950, have been described as random graphs.

According to Cohen and Havlin (2010), complex networks describe many systems in nature and society. Examples often cited include cells in the human body, in which chemical compounds are connected through chemical reactions. Another example is the Internet, in which computers and routers connected by cables form computer networks.

The growing study of complex networks began to question whether these are derived from random networks. However, it was necessary to quantify some principles of organization of these networks, which differentiate them from the characteristics of random networks.

2.3 About Scientometrics

Citations are also relationships over time between previous publications of their references and future citations of the work in question. Eugene Garfield's seminal work in 1950 - which identified the importance of citations and conceived the Science Citation Index (SCI) - culminated in the creation of Information, ISI, which quickly established itself as a massive citation collection database. (Garfield, 1955; Garfield, 1979) Interestingly, the initial focus was on helping researchers explore the literature and not on quantitative research evaluation.

It turns out that citations are also beneficial as research evaluation metrics, enabling authors and their work to be discovered. Thus, the SCI became part of the Social Sciences Citation Index (SSCI, in 1973). Furthermore, five years later, they were added to the Arts & Humanities Citation Index (A&HCI), acquired by Thomson Corporation, which developed Web of Science as part of its Web of Knowledge platform. In 2013, the SCI covered 8,539 journals, the SSCI 3,080 journals, and the A&HCI approximately 1,700 journals. As early as 1973, several researchers and research institutes recognized the relevance of the SCI to empirical research on the international practice of scientific activity a.

The physicist and historian of Science Derek de Solla Price (1965 and 2011) realized the importance of networks of articles and authors. Also, they began to analyze scientometric processes, giving rise to the idea of cumulative advantage (Price, 1976), a version of "success to the successful" (Senge, 1990) or "Success begets success (SBS)," also known as the Matthew Effect (Merton, 1968; Merton, 1988).

Price identified some of the main problems that scientometrics would address: mapping the "Invisible Faculties" (Crane, 1972) informally linking highly cited researchers at the frontiers of research (cf co-authorship networks and co-citation analysis (Marshakova, 1973; Small, 1973)): study the links between productivity and quality in which the most productive are often the most cited (according to the h-index), and investigating citation practices in different fields (as per standardization).

In 1968 and 1988, Robert K. Merton, a leading sociologist, was one of those who explored many of these new scientometric approaches. Scientometrics also gained importance as a branch of research with greater autonomy with the creation of (i) Scientometrics in 1978, (ii) a research center at the Hungarian Academy of Sciences, and (iii) events and academic associations focused on the theme.

3. Methodology

3.1 Data

The paper's data correspond to the period from 1896 to 2006. The sample was based on bibliographic references considered relevant and cited by Gonçalo Fonseca and Peter L. Bernstein. Gonçalo Fonseca is a historian of economic thought who maintains the main international website in the field: *The History of Economic Thought* from the *Institute for New Economic Thinking* (Note 1). Peter L. Bernstein is perhaps the best-known financial historian. He wrote the bestseller "Dare to the Gods: The Fascinating History of Risk" and the "History of Capital Markets." We base ourselves on the literature of this second book, in which the author selected the main works that consolidated what he called "capital ideas." These ideas are the most influential theoretical concepts and applied studies that have consolidated the modern practices of participants in financial markets, in addition to basing management practices and financial decisions on the capital, derivatives, and insurance markets, among others.

Various articles use computationally intensive data collection processes (web scrapping). However, we decided to follow a strategy of taking advantage of the expertise of two specialists in the area of the history of economic-financial thought in this paper. We supported the critical view of both and chose to work with a smaller sample but more precise, specific, and carefully selected by these two specialists. We assume that a smaller sample collected from the selection of two specialists in the literature on a specific topic can be more representative than a large sample in which relevant authors are mixed with many others without expression in terms of prestige, academics, and impact on scientific literature.

3.2. Complex Network Metrics

This work uses a directed network (digraph, digraph, or directed graph). It is analyzed at the level of agents (using a measure of centrality, the weighted degree) and at the level of groups, as the identification of communities (clusters). Such identification is made with the modularity optimization process.

3.2.1 Degree Centrality

The degree of a node is characterized by the number of edges connected to it. According to Newman (2003a and 2003b), the degree of distribution is divided into in-degree and out-degree in directed graphs, which represent respectively the number of citations received by a node and the number of citations executed by a node. The entry degree of a node can indicate its importance and imply the possibility of being cited in the future. However, some metrics, such as eigenvector centrality, can characterize this importance differently.

In random networks, as the edges are inserted randomly, most nodes have the same number of edges, which is close to the average degree of the network. The distribution of degrees in the random lattice follows the Poisson distribution.

For complex networks, Albert and Barabási (2002) found that the degree distribution follows a power-law tail distribution different from random networks. An example of a network with this type of distribution is the World Wide Web[12] and the Internet[13].

3.2.2 Modularity

One of the unique characteristics of social networks is that they have a community structure. Usually, this property emerges as a consequence of the global and local heterogeneity of the distribution of edges in a graph, so in this type of network, it is possible to find high concentrations of edges in specific regions and low concentrations of edges between these regions.

Communities, or clusters, are groups of densely connected vertices with sparse links. According to Newman and Girvan (2004), there are two main lines of investigation in discovering network communities. The first originated in Computer Science and is known as graph partitioning. At the same time, the second was essentially developed by sociologists and is usually referred to as blockmodeling, hierarchical clustering, or community structure detection. The community detection algorithms' basic process is based on dividing the original graph into a set of disjoint subgraphs by optimizing a given objective function. Both approaches aim to discover groups of related vertices and, if possible, define their hierarchical organization based on information provided by the network topology. It is usually accomplished by iteratively removing bridging edges connecting groups of vertices, as suggested by Girvan and Newman (2002).

In real life, it is possible to find various examples of cohesive groups or communities. Society is an environment rich in finding communities, as people naturally tend to form groups. These groups can be families, circles of friends, religious or work groups, cities, and nations, among others. If we also consider groups formed by companies or consumers of a given product, it is possible to identify communities relevant to the area of Economics and Management.

The importance of studying these communities is intuitive in domains such as ARS.

Fortunato (2010) stated that the structural assumption of nodes in each network community could help the identification of central actors associated with stability and group control functions. The intermediate actors are located on the borders of communities and play a key role in disseminating and exchanging information. The intermediate actors create bridges between communities.

Hierarchical clustering is a class of methods for detecting clusters or groups. Hierarchical algorithms generate structures of groups inserted into larger groups that, in turn, are inserted into even larger groups, which are represented by dendrograms that show the multilevel structure of the network. These methods effectively solve group analysis problems and similar problems such as graph partitioning and community identification.

Hierarchical grouping is also quite intuitive and is based on the definition of similarity. First, it is necessary to choose a measure of similarity (or dissimilarity) to assess how similar two nodes are, according to a given global or local property. Then, the similarity matrix between all pairs of nodes must be calculated, regardless of whether these nodes are connected. Then, it is necessary to select a method to group the nodes: the agglomerative methods, which focus on the denser regions of the network instead of focusing on the connections at the edges of the network (Note 2), or divisive methods, which focus on identifying and removing links that connect densely connected regions to the network, especially bridges and local bridges (Easley and Kleinberg, 2010) (Note 3).

Depending on the choice, a distance measure is selected to calculate the similarity between groups (Note 4). The final result of this process is a dendrogram that illustrates the organization of nodes returned by the hierarchical Algorithm. One strategy to select the best methods is calculating the community's modularity value (Newman and Girvan, 2004) and selecting the number that maximizes this function.

Modularity (class) optimization is another method used to detect network communities. Q modularity is a quality function that evaluates and measures the importance of a given network partition into communities. This function is used to compare the quality of partitions and as an objective function in optimization problems. According to Newman (2006), modularity is represented by the normalized difference between the number of edges observed within each group of nodes in the network and the number of edges observed within the same group in the network for randomly generated edges. The modularity Q is calculated as follows:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

The m indicates the number of edges; k_i and k_j respectively represents the degree of the vertices i and j; A_{ij} is the input of the adjacency matrix that indicates the number of connections established between the vertices *i* and *j*; $\frac{k_i k_j}{2m}$ represents the expected number of edges that should exist between the pair of vertices *(i, j)*; c_i and c_j denote the groups to which the vertices *i* and *j* belong; and $\delta(c_i, c_j)$ represents the Kronecker delta.

Modularity *Q* can take on positive and negative values. If $Q > 0$, a networked community structure is possible. If *Q* is a large positive number, then the respective partition is more likely to reflect the proper community structure. According to Clauset, Newman, and Moore (2004), modularity that assumes values greater than or equal to 0.3 is a good indicator of a meaningful community in the network.

A common feature in many social networks is the detection of community structure, which consists of dividing network nodes into groups with dense internal and sparse external connections. The study of community structures in social networks relates to graph partitioning in graph theory and computer science and hierarchical clustering in sociology.

According to Jain, Murty, and Flynn (1999), the definition of clustering is the unattended recognition of patterns in data. This technique is beneficial for analyzing patterns, clusters, decision-making, machine learning, and especially exploring relationships between nodes in a network.

Modularity is a metric that can measure the quality of the graph's division, dividing the network's various nodes into groups. By grouping the nodes, it is possible to verify them with a certain degree of similarity, although they are not directly connected. Modularity is part of community detection research, which is associated with graph theory and hierarchical clustering. They use methods to divide the network into subgraphs representing each existing community in the network. The network is considered random when the modularity value is close to zero. The community has a strong structure when the value is close to one. According to Girvan and Newman (2002), in practice, the values are between 0.3 and 0.7.

4. Results

To analyze the results, we used two metrics of complex networks (weighted degrees and modularity), the scores and the IDEAS/RePEc ranking. The weighted degrees are in descending order. Moreover, it is observed that there is a prominence of cluster number 47, composed by the names in bold in Table 1 and described in detail in Table 2, namely: Daniel Kahneman, Luigi Zingales, Robert Kosowski, Allan Timmermann, Russell Wermers, and Halber White. It is because of a widely cited article involving the last four authors and because Kahneman and Zingales have important scores and prominent positions in the IDEAS/RePEc ranking.

In Table 1, we can also assess the importance of the seminal articles on the Black-Scholes Option Pricing Model and the Modigliani-Miller Theorem, which will be discussed later.

Still, Table 1 shows financial economists who stand out in their research activities as authors or co-authors. Table 1 presents the top 175. Nevertheless, we could highlight that, based on the weighted degrees, the 15 with the most relevant production are, in descending order: Daniel Kahneman, Robert Kosowski, Luigi Zingales, Stephen A. Ross, Eugene Fama, Merton Miller, Fisher Black, Franco Modigliani, Myron Scholes, Sanford J. Grossman, Oliver D. Hart, Kenneth French, Robert Lucas Jr., Nicholas Cox, and Edward C. Prescott.

In the last two columns of Table 1, we record the IDEAS/RePEc scores and ranking. Unfortunately, such data are not available for all 175 researchers. We found a negative correlation (-0.2204) between the weighted degrees and the scores and positions in the rankings of economists with both metrics. The result is evident. The lower the score of a researcher, the higher his position in that ranking. Thus, the lower the score, the higher the weighted grade and the higher the number of citations the researcher received on Google Scholar.

Table 1. Weighted degrees, modularity classes, score, and ranking

Id	Researchers in Financial Economics	Weighted degrees	Modularity classes	Score IDEAS/RePEc	Ranking IDEAS/RePEc
107	Daniel Kahneman	150976	47	145.59	123
113	Robert Kosowski	70504	47		
112	Luigi Zingales	67295	47	87.49	71
17	Stephen A. Ross	62409	8	234.81	202
43	Eugene Fama	60798	32	16.09	9
28	Merton Miller	59714	11	828.74	756
18	Fisher Black	58225	11		
29	Franco Modigliani	52000	11	776.74	700
19	Myron Scholes	51637	11	727.17	650
32	Sanford J. Grossman	41023	12		
21	Oliver D. Hart	33929	12	158.16	140
83	Kenneth R. French	32847	32	50.5	38
44	Robert E. Lucas Jr.	31550	100	22.81	12
24	Nicholas Cox	28223	8	48.31	34
211	Edward C.Prescott	20116	100	60.69	45
15	Richard Roll	17217	8	463.22	411
109	Tversky	16641	47		
214	J. MacBeth	16588	32		
27	Ingersoll Jr.	14210	8		11
63	John Y. Campbell	12859	23	20.04	11
160	Nancy L. Stokey	10849	100	896.42	818
33	Joseph E. Stiglitz	10525	12	6.16	4
25	Ariel Rubinstein	9414	8	300.42	277
35	Douglas Gale	8631	95	332.19	306
46	Robert J. Shiller	8369	23	89.72	75
6	Jacob Marschak	7566	2		

50	Brad M. Barber	7223	18	751.25	672
77	Andrei Shleifer	7132	71	3.19	1
189	Roy George D. Allen	7115	95		
237	Daniel Orr	7019	12		
146	Robert W. Vishny	6094	71	42.92	29
22	David M. Kreps	5598	102	482.13	432
34	Paul Milgrom	5306	100	92.71	81
222	J.Michael Harrison	5255	102		
53	Terrance Odean	5195	18	862.14	788
30	Michael C. Jensen	5001	11	61.07	46
12	William J. Baumol	4404	71	191.52	170
42	Hendrik S. Houthakker	4097	98		
54	Richard H. Thaler	3578	75	100.14	84
8	Benjamin Graham	3571	4		
201	David Dodd	3571	4		
220	John Roberts	3124	100	525.8	466
11	James Tobin	3044	6	298.21	272
151	Eric J. Johnson	3027	75		
76	James Bradford DeLong	2781	71	869.85	795
206	William C. Brainard	2565	6		
219	Martin F. Hellwig	2525	95	518.48	456
16	Roy Radner	2391	2		
64	Jens Hilscher	2314	23		
65	Peter G. Szilagyi	2314	23		
57	Robert Litterman	2165	11	2700.13	2589
39	Oskar Morgenstern	1701	77		
217	Stephen P. Magee	1534	98		
10	Harry M. Markowitz	1495	44	1115.9	1029
213	Marshall E. Blume	1478	32		
215	Sigbert J. Prais	1443	98		
210	Leonard A. Rapping	1438	100		
66	Andrew W. Lo	1429	56	439.59	388
127	A. Craig MacKinlay	1429	56		
174	Kenneth J. Arrow	1429	2	235.81	205
198	Theodore Harris	1429	2		
138	Raghavendra Rau	1356	67		
212	Thomas J. Sargent	1329	100	43.26	30
73	Daniel Cohen	1258	11	1639.34	1536
197	W.H. Andrews	1239	2		
20	Robert C. Merton	1211	9	86.33	70
59	Markus K. Brunnermeier	1200	20	313.69	294
60	Stefan Nagel	1200	20	704.12	625
4	John R. Hicks	1171	95	2583.19	2466

114	Allan Timmermann	1159	47	476.79	424
115	Russell Wermers	1159	47		
116	Halbert White	1159	47	105.7	93
203	Haim Levy	1134	44		
216	Lester D. Taylor	1120	98		
156	John Von Neumann	1117	77		
155	Jack Treynor	1081	11		
51	Yi-Tsung Lee	1014	18		
52	Yu-Jane Liu	1014	18		
70	George R.G. Clarke	894	26		
224	Frank J. Fabozzi	806	11	2249.14	2142
238	Charles W. Upton	768	11		
5	Nicholas Kaldor	767	1		
194	James A. Mirrlees	764	1	752.31	674
200	H.D. Block	717	2		
139	Valentin Dimitrov	678	67		
140	Melinda Cooper	678	67		
209	Terry Marsh	657	9		
232	Wiliiam G. Bowen	634	71		
159	Clive W. J. Granger	584	77	71.41	58
85	Xavier Gabaix	578	34	190.27	169
223	Stanley Fischer	571	11		
40	Paul A. Samuelson	552	9	119.91	106
110	Hyeonsoo Kim	536	47		
23	Darrell Duffie	535	101	185.98	164
221	Wayne Shafer	535	101		
38	Alfred Cowles	481	99		
218	Homer Jones	481	99		
71	Harindra de Silva	447	26		
72	Steven Thorley	447	26		
231	T.Fabian	436	71		
227	Edward A. Ide	412	71		
103	Bruce I. Jacobs	366	44		
229	Philip H. Dybvig	352	8	910.57	832
236	Dwight M. Jaffee	326	12		
207	Robert M. Solow	320	6	266.83	235
153	Peter Temin	292	76		
154	Hans-Joachim Voth	292	76	1194.83	1108
86	Arvind Krishnamurty	289	34		
87	Olivier Vigneron	289	34		
233	Richard Emeric Quandt	283	71		
148	Charles Smithson	282	73		
149	Betty Simkins	282	73		

152	Jack L. Knetsch	271	75	1769.44	1666
111	Adair Morse	268	47		
235	Richard A. Posner	261	11	1917.79	1802
234	S.A. Batey Blackman	258	71		
230	Michael B. Connolly	258	8		
93	Richard Grinold	257	39		
94	Ronald Kahn	257	39		
228	R.E. Gomory	232	71		
89	Amit Goyal	232	36	1864.93	1743
90	Sumil Warhal	232	36		
199	D. Davidson	206	2		
104	Kenneth Levy	183	44		
202	Laurie A. Goodman	179	44		
225	Gary S. Becker	147	71	22.96	14
192	S. Hollander	143	95		
47	Jeffrey J. Anderson	142	16		
48	Vernon L. Smith	142	16	302.49	283
3	John Maynard Keynes	140	1		
45	Burton Gordon Malkiel	138	71		
106	Jeeman Jung	138	23		
196	H. Makower	127	2		
67	Nicholas Chan	107	24		
68	Mila Getmansky	107	24		
101	Martin Leibowitz	96	52		
121	Owen A. Lamont	92	51	1388.1	1295
122	Jeremy C. Stein	92	51	112.63	102
239	Christopher L. Culp	91	11		
240	Andrea M. P. Neves	91	11		
31	Jacques H. Drèze	79	11	710.56	632
204	Harold W. Watts	73	6		
145	Eduardo Schwarz	66	70		
226	M.H. Preston	52	71		
2	Irving Fisher	52	6		
123	Anthony Bova	49	52		
100	Sidney Homer	47	52		
96	W. V. Harlow	41	41		
97	K. Brown	41	41		
99	Joanne Hill	33	70		
144	Thomas Schneeweis	33	70		
81	Elroy Dimson	28	31	3232.23	3096
82	Paul Marsh	28	31		
195	Emil Lederer	28	2		
178	E.L. Fisk	24	6		

183	H.D. Henderson	19	1		
190	Ursula K.Hicks	17	95		
184	A. Salter	17	1		
185	Josiah C. Stamp	17	1		
186	Basil Blackett	17	1		
187	Henry Clay	17	1		
188	William H. Beveridge	17	1		
177	Alban William Phillips	14	6		
180	E. Cannan	11	1		
181	Charles Addis	11	1		
182	Alfred Milner	11	1		
117	Mark Kritzman	10	48		
118	Lee Thomas	10	48		
55	Shlomo Benartzi	9	75		
205	F. Trenery Dolbear	7	6		
143	William F. Sharpe	6	44	964.87	883
179	E. F Robbins	4	6		
74	Marvin Damsma	3	27		
75	Gregory Williamson	3	27		
193	Arjo Klamer	2	95		
56	Peter L. Bernstein	2	47		

Source: elaborated by authors.

In Table 2, we have the seven main clusters of the sample with the respective researchers and their lines of research. The definition of the most relevant clusters was based on the average weighted degrees of the articles in co-authorship most cited by Google Scholar until 05/17/2021.

The first cluster in descending order of importance, whose modularity class is 47, brings together two researchers highlighted in bold in Table 1: Daniel Kahneman and Luigi Zingales. The former is the main reference in behavioral finance and perspective theory. The second has much-cited works involving capital structure, financial development, and the relationships between financial markets and economic growth. It is also worth mentioning that another important co-author on behavioral finance and perspective theory (prospect theory, together with Kahneman) is also present in this cluster: Amos Tversky.

Table 2. Clusters, average of weighted degrees, researchers, and lines of research

Cluster	Average of weighted degrees	Researchers	Research lines
47	41,788.57	Daniel Kahneman Robert Kosowski Luigi Zingales Allan Timmermann Russell Wermers Amos Tversky Halbert White Adair Morse	Behavioral finance and perspective theory (Kahneman); performance analysis of mutual fund managers in the long term (Kosowski, Timmerman, Wermers, and White); capital structure, financial development, finance, and economic growth (Zingales; Morse).
8	30,514.75	Stephen A. Ross Nicholas Cox Richard Roll Amos Tversky Ingersoll Jr. Ariel Rubinstein Philip H. Dybvig Michael B. Connolly	Theory of the term structure of the interest rate (Cox, Ingersoll, and Ross); option pricing (Cox, Ross, and Rubinstein); Behavioral Finance and perspective theory (Tversky); Portfolio theory and asset pricing (Roll and Ross); Exchange, interest and financial assets in emerging countries (Connolly); Banking crises (Dybvig) and portfolio theory (Dybvig and Ross).
32	27,927.75	Eugene Fama J. MacBeth Marshall E. Blume Kenneth R. French	Efficient markets hypothesis (Fama and French); Method for calculating betas and analyzing the debt quality of U.S. companies (Blume); Fame-MacBeth Regression (Fame and MacBeth).
12	23,124.00	Sanford J. Grossman Oliver D. Hart Joseph E. Stiglitz Daniel Orr	Theory of vertical and horizontal integration (Grossman and Hart); Asymmetric information theory and efficient markets (Grossman, Stiglitz); Competitive balance in the stock market (Grossman and Hart); Demand for currency by companies (Miller and Orr).
11	16,677.64	Merton Miller Fisher Black Franco Modigliani Myron Scholes Robert Litterman Daniel Cohen Michael C. Jensen Jack Treynor Frank J. Fabozzi Charles W. Upton Stanley Fischer Christopher L. Culp Andrea M. P. Neves Jacques H. Drèze	Black-Scholes Model; Modigliani-Miller Theorem; Asset allocation, fixed income bond returns and global portfolio optimization (Litterman and Black); Portfolio theory (Black, Jensen and Scholes, Treynor); Financial institutions and financial instruments (Modigliani and Fabozzi); Financial Services Cost and Hotelling's Valuation Principle Test (Miller and Upton); Effects and costs of inflation on assets (Fischer and Modigliani); Banking regulation (Black, Miller, and Posner); Var - Value at Risk (Miller, Culp, and Neves); Stochastic financial planning (Miller and Drèze).

		Richard A. Posner	
100	9,214.00	Robert E. Lucas Jr. Edward C.Prescott Nancy L. Stokey Paul Milgrom John Roberts Leonard A. Rapping Thomas J. Sargent	Inflation, real wages and employment (Lucas and Rapping); Investment under uncertainty and Equilibrium and Unemployment (Lucas and Prescott); Recursive Methods in Dynamic Economics, Currency and Interest in a Cash-in-Advance Model, Optimal Monetary and Fiscal Policy and Optimal Growth (Lucas and Stokey); New-Classical Macroeconomics (Lucas and Sargent); Price Signaling (Milgrom and Roberts); Information Economics (Milgrom and Stokey).
23	6,464.00	John Y. Campbell Robert J. Shiller Jens Hilscher Peter G. Szilagyi	Trading volume and serial correlation in equity returns (Campbell and Grossman); Determinants of company bankruptcy and share price of insolvent companies (Campbell, Hilscher, and Szilagyi); Financial markets are macro-inefficient, but micro-efficient - Samuelson's phrase (Jung, Jeeman and Shiller); Expectations about future dividends, valuation ratios, and long-term capital markets; Investments in large companies; Interest rate movements and increase in earnings (Shiller).

Source: elaborated by authors.

The second most important cluster, modularity class 8, aggregates works by Stephen A. Ross, Phillip Dybvig, and others. The most important topics they researched were: the interest rate term structure theory (Cox, Ingersoll, and Ross); option pricing (Cox, Ross, and Rubinstein); Portfolio theory and asset pricing (Roll and Ross); Exchange, interest, and financial assets in emerging countries (Connolly); Banking crises (Dybvig) and portfolio theory (Dybvig and Ross).

The third cluster is influenced by works on the efficient markets hypothesis, especially by the Fama-French and Fama-MacBeth models. The other four clusters are described in detail in Table 2. We can observe each cluster's respective weighted average degrees, modularity classes, researchers, and research lines in it.

Finally, we describe a complex network, in Figure 4, with the clusters separated by modularity classes. We used the proportional Yifan Hu distribution as a layout algorithm. Nevertheless, before we comment on it, let us define the Yifan Hu algorithm that originated it. In the words of Yifan Hu himself (2006):

> "We propose a graph drawing algorithm that is efficient and of high quality. This Algorithm combines a multilevel approach, which effectively overcomes local minima, with Barne and Hut's eight-tree technique, which efficiently approximates short-range and long-range forces. Our numerical results show that the Algorithm is comparable in speed to Walshaw's highly efficient multilevel graph drawing algorithm and even yields better results for some of the difficult problems. In addition, an adaptive attenuation scheme for force-driven algorithms and a general repulsive force model are proposed." (Yifan Hu, 2006).

Therefore, the Yifan Hu Proportional algorithm is similar to the Yifan Hu algorithm. The difference is that the first establishes a proportional displacement to distribute the vertices in the area of the graph. In terms of calculation speed and accuracy, there is not much difference between the two.

We also used 25 different colors to describe the main clusters categorized by their respective weighted degrees. The most peripheral clusters in the network that received the gray color are the least relevant. On the other hand, those that received colors and are closer to the center of the network are the most representative regarding the number of citations of articles written in co-authorship in the financial economics literature.

An important detail in the network in Figure 4 is that thicker edges connecting the vertices (agents or co-authors) represent the importance of co-authorship in terms of citations in Google Scholar. Thus, we can observe that these strong links reveal the prestige of co-authorships among researchers who have been noted for their contributions (summarized here) to the financial literature. Among which we can mention:

- Kahneman and Tversky: Prospect Theory - explains how people make decisions about alternatives that involve risks, in which the probabilities of outcomes are uncertain. Both authors argue that people make decisions more influenced by the probable values of losses and gains than by examining the possible final result. They also suggested that a loss can have a greater emotional impact on an individual than an equivalent gain. Therefore, it is likely that a person will work harder to avoid a loss than to try to gain a gain. Some behaviors observed in the economy, such as changes in risk aversion, can be explained by Prospect Theory. Investors can sell appreciated assets, making gains quickly, while tending to hold depreciated assets. It can reduce gains and increase losses on investments.

- Kahneman and Zingales: Kim, Morse, and Zingales (2006) mention the article by Kahneman and Tversky as the second most cited in the literature for 35 years, from 1970 to 2005. The sample only gathered articles with more than 500 citations.

- Modigliani and Miller: Modigliani-Miller theorem, which is also called the capital structure irrelevance principle, as it argues that in an efficient market without taxes, bankruptcy costs, agency costs, and asymmetric information, the value of a company is not influenced by its capital structure, that is, how this company is financed. Therefore, the company's value does not depend on its dividend policy or its decision to increase capital by issuing shares or selling debt.

- Grossman and Hart (G.H. model) can be interpreted as a theory of markets rather than a theory of the firm. For Holmström (2016), the G.H. model is the first theory that explains why markets are critical to organizational choice and that the virtue of market transactions stems from the retention power conferred by ownership. In particular, G.H. provides a new answer to the Williamson puzzle of selected intervention. That answers why integrated companies are not always successful in replicating a market outcome.

- Cox, Ingersson, and Ross: Cox-Ingersson-Ross model describes the evolution of interest rates. It is a "one-factor model" (short rate model), as it presents interest rate movements as being driven by only a single source of market risk. This model is adopted as a valuation method for interest rate derivatives. It was introduced in 1985 by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross as an extension of the Vasicek model.

- Black and Scholes: Black-Scholes formula - prices the theoretical value of a call or put option using six variables as references: volatility, type of option, price of the underlying stock, time, exercise price, and risk-free rate.

- Lucas and Prescott – the strong link (thick edge or "heavy") refers to the classic article "Investment under Uncertainty," published in Econometrica in September 1971. The article describes the behavior of investment, product, and price time series in a competitive industry with stochastic demand. They show that the equilibrium for the industry is obtained by solving a dynamic programming problem (maximizing consumer surplus). Moreover, after solving this question, they determine the characteristics of equilibrium trajectories. This article received 19,189 citations on Google Scholar as of 5/17/2021.

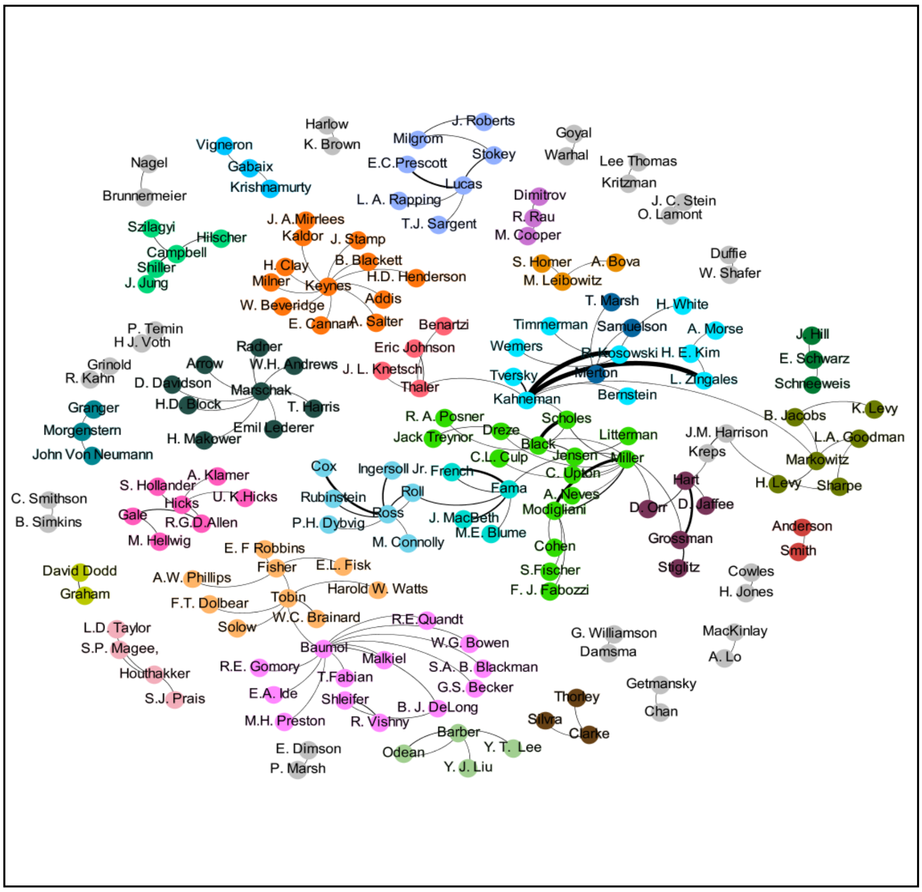

Figure 4. Leading Finance Research Communities

Source: elaborated by authors.

5. Final Remarks

The work tried to answer three questions related to economic-financial scientometrics: (i) how to detect the most cited authors and co-authors in a sample of the most influential works in the literature on financial economics? (ii) How to define this sample's most relevant co-authorship groups? (iii) How to elaborate a complex network capable of reliably describing the links between these clusters of authors and co-authors, highlighting the most significant lines of research in the financial economics literature from 1896 to 2006.

We use two metrics of complex networks to answer these questions: the weighted degree and the modularity class. Then, we compare them with the IDEAS/RePEc scores and ranking. To describe the network, we used Yifan Hu's proportional layout algorithm. The database was collected from two sources: the main international website for the History of Economic Thought: the History of Economic Thought from the Institute for New Economic Thinking, a page maintained by the historian of economic thought Gonçalo Fonseca; and the references described by financial historian Peter L. Bernstein in his seminal book "History of the Capital Market."

We show that - in a table with descending order of weighted degrees, which detects the most cited authors and co-authors on Google Scholar up to 05/17/2021 - there is a preponderance of cluster number 47. It is composed by Daniel Kahneman, Luigi Zingales, Robert Kosowski, Amos Tversky, Allan Timmermann, Russell Wermers and Halber White. Because of a widely cited article involving the last four authors and because Kahneman and Tversky have written many seminal articles in pairs and Kahneman and Zingales have important scores and prominent positions in the IDEAS/RePEc ranking.

In this same list, we measure and rank the importance of seminal articles on the Black-Scholes Option Pricing Model and the Modigliani-Miller Theorem. Furthermore, in quantitative terms, we compiled a ranking of 175 financial economists who stood out in their research activities as authors or co-authors. The top 15 in this ranking were Daniel Kahneman, Robert Kosowski, Luigi Zingales, Stephen A. Ross, Eugene Fama, Merton Miller, Fisher Black, Franco Modigliani, Myron Scholes, Sanford J. Grossman, Oliver D. Hart, Kenneth French, Robert Lucas Jr., Nicholas Cox, and Edward C. Prescott.

We verified a negative correlation (-0.2204) between the weighted degrees and the scores and positions in the rankings of economists with both metrics. Such a result is intuitive, given that the lower a researcher's score, the higher his position in that ranking. Thus, the lower the score, the higher the weighted grade and the higher the number of citations the researcher received on Google Scholar.

We created a complex network in which we separated the main clusters or groups of researchers in Financial Economics with 25 different colors. In this network, the thicker edges connecting the vertices (authors or co-authors) represent the importance of co-authorship in terms of citations in Google Scholar. With this, we show that the strong links reveal the prestige of co-authorships among researchers who have been noted for their contributions (summarized here) to the financial literature. Among them, we can mention Kahneman and

Tversky (Prospect Theory); Kahneman and Zingales: Kim, Morse, and Zingales (2006) mention the article by Kahneman and Tversky as the second most cited in the literature for 35 years, from 1970 to 2005.

The sample only gathered articles with more than 500 citations. Modigliani and Miller (Modigliani-Miller Theorem); Grossman and Hart (G.H. Model can be interpreted as a theory of markets rather than a theory of the company); Cox, Ingersson and Ross (Cox-Ingersson-Ross Model, which describes the evolution of interest rates); Black and Scholes (Black-Scholes Formula - prices the theoretical value of a call or put option using six variables as references).

Lastly, Lucas and Prescott – the strong link (thick edge or "heavy") refers to the classic article "Investment under Uncertainty," published in Econometrica in September 1971. The paper describes the behavior of investment, product, and price time series in a competitive industry with stochastic demand. They show that the equilibrium for the industry is obtained by solving a dynamic programming problem (maximizing consumer surplus). Furthermore, after solving this question, they determine the characteristics of equilibrium trajectories. This article received 19,189 citations on Google Scholar as of 5/17/2021.

As a suggestion for future research, the most prestigious areas of the entire economic literature could be studied, not only financial economics, which was the focus of this work.

References

Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. Reviews of modern physics, 74(1), 47.

Albert, R., Jeong, H., & Barabási, A. L. (1999). Diameter of the world-wide web. Nature, 401(6749), 130-131.

Amon, J., & Hornik, K. (2022). Is it all bafflegab? – Linguistic and meta characteristics of research articles in prestigious economics journals. Journal of Informetrics, 16(2), 101284.

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286(5439), 509-512.

Barabási, A. L., & Bonabeau, E. (2003). Scale-free networks. Scientific American, 288(5), 60-69.

Bernstein, P. L. (2008). História do Mercado de Capitais. Rio de Janeiro: Elsevier.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, 70(6), 066111.

Cohen, R., & Havlin, S. (2010). *Complex Networks: Structure, Robustness and Function*. Cambridge, Massachussets: Cambridge University Press.

Crane, D. (1972). *Invisible Colleges*. Chicago: University of Chicago Press.

Easley, D., & Kleinberg, J. (2012). Networks, crowds, and markets: Reasoning about a highly connected world. Significance, 9(1), 43-44.

Erdős, P., & Rényi, A. (1959). On Random Graphs. *Publicationes Mathematicae*, 6, 290-297.

Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*.

Fortunato, S. (2010). Community detection in graphs. *Physics Report, 486*(3-5), 75-174.

Franceschet, M. (2010). The difference between popularity and prestige in the sciences and in the social sciences: A bibliometric analysis. *Journal of Informetrics*, *4*(1), 55-63.

Galiani, S., & GÁlvez, R. H. (2019). An empirical approach based on quantile regression for estimating citation ageing. *Journal of Informetrics*, *13*(2), 738-750.

Garfield, E. (2008). *HistCite: Bibliometric Analysis and Visualization Software (Version 8.5.26).* Bala Cynwyd, PA: HistCite Software LLC. Retrieved July 15, 2008, from http://www.histcite.com/

Garfield, E., Cronin, B., & Atkins, H. B. (2000). *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield*. Medford, N.J.: Information Today.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. Proceedings of the national academy of sciences, 99(12), 7821-7826.

Girvan, M., & Newman, M. E. J. (2003). Finding and evaluating community structure in networks. *Phys. Rev. E,* 69(026113).

Hu, Y. (2006). Efficient, high-quality force-directed graph drawing. *The Mathematica Journal*, *10*(1).

Halevi, G., Moed, H., & Bar-Ilanc, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation — Review of the Literature. *Journal of Informetrics*, *11*(3), 823-834.

History of Economic Thought Website. (2021). Website on the history of economic thought maintained and written by FONSECA, G. and supported by the Institute for New Economic Thinking. Retrieved from http://www.hetwebsite.net/het/home.htm

Holmström, B. (2016). *Grossman-Hart (1986) as a Theory of Markets*. https://doi.org/10.1093/acprof:oso/9780199826223.003.0002

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data Clustering: A Review. *ACM Comput. Surv.*, *31*(3), 264-323.

Kossinets, G., & Watts, D. J. (1995). Neural network ensembles, cross validation and active learning. In Freeman, L. C., Thompson, C. R., & Kochen, M. (Eds.), *Advances in Neural Information Processing Systems* (1989, v. 7, p. 231-238). The small world.

Macohin, A., & Serbena, C. A. (2015). Legal precedents analysis through complex networks. In *ICAIL Multilingual Workshop on A.I. & Law Research.*

Marshakova, I. V. (1973). Co-Citation in Scientific Literature: A New Measure of the Relationship Between Publications. *Scientific and Technical Information Serial of VINITI, 6*, 3-8.

Merton, R. (1968). The Matthew effect in Science. *Science*, *159*(3810), 56-63.

Merton, R. (1988). The Matthew effect in Science II: Cumulative advantage and the symbolism of intellectual property. *ISIS*, *79*, 606-623.

Mingers, J., & Leyersdorff, L. (2015). A review of theory and practice in Scientometrics. *European Journal of Operational Research*, *246*(1), 1-19.

Newman, M. E. J. (2003a). Mixing patterns in networks. *Physical Review E*, *67*(2): 026126.

Newman, M. E. J. (2003b). The structure and function of complex networks. *SIAM Review, 45*(23), 167-228.

Newman, M. E. J. (2004). Fast Algorithm for detecting community structure in networks. *Physical review E*, *69*(6).

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America, 103*(23), 8577-8528.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E, 69*(2), 026113.

Pons, P., & Latapy, M. (2006). Computing commuties in large networks using random networks. *Journal of Graph Algorithms and Applications, 10*(2), 191-2018.

Price, D. J. S. (1965). Networks of Scientific Papers -The pattern of bibliographic references indicates the nature of the research front. *Science*, *149*(3683), 157-162.

Price, D. J. S. (2011). Networks of scientific papers, *The Structure and Dynamics of Networks*, 149-154.

Seiler, C., & Wohlrabe, K. (2012). Ranking economists on the basis of many indicators: An alternative approach using RePEc data. *Journal of Informetrics*, *6*(3), 389-402.

Seiler, C., & Wohlrabe, K. (2014). How robust are journal rankings based on the impact factor? Evidence from the economic sciences. *Journal of Informetrics*, *8*(4), 904-911.

Small, H. (1973). Co-Citation in Scientific Literature: A New Measure of the Relationship Between Publications. *JASIS*, *24*, 265-269.

Tol, R. S. (2013). The Matthew effect for cohorts of economists. *Journal of Informetrics*, *7*(2), 522-527.

Wu, D., Li, J., Lu, X., & Li, J. (2018). Journal editorship index for assessing the scholarly impact of academic institutions: An empirical analysis in the field of economics. *Journal of Informetrics*, *12*(2), 448-460.

Notes

Note 1. https://www.ineteconomics.org/education/materials/history-of-economic-thought-website

Note 2. The Walktrap algorithm (developed by Pons and Latapy, 2005) is an example of this type of method.

Note 3. The well-known algorithm by Girvan and Newman (2002) is a well-cited example of this method.

Note 4. Some examples of this measure: the single linkage (or nearest neighbor), the complete linkage (or furthest neighbor) and Ward's method.