The Nobel Prize in Economics: individual or collective merits?

José Alberto Molina[1,4,6*] , David Iñiguez[3,4,5], Gonzalo Ruiz[4,5], Alfonso Tarancón[2,4,5]

1. Departamento de Análisis Económico, Universidad de Zaragoza, Zaragoza, Spain
2. Departamento de Física Teórica, Universidad de Zaragoza, Zaragoza, Spain
3. Fundación ARAID, Diputación General de Aragón, Zaragoza, Spain
4. Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Zaragoza, Spain
5. Kampal Data Solutions S.L., Zaragoza, Spain

Abstract

We analyse the research production of Nobel laureates in Economics, employing the JCR Impact Factor (IF) of their publications. We associate this production indicator with the level of collaboration established with other authors, using Complex Networks techniques applied to the co-authorship networks. We study both individual and collaborative behaviours, and how th

collaborators, who have also contributed to the discovery or results being rewarded.

Network methods, initially derived from Physics and Computational Sciences, have been increasingly applied to study scientific output in different fields of research (Ecology [14], History and Humanities [15-16], Medicine and Biology [17-18], Nanotechnology [19], Nanoscience, Pharmacology and Statistics [20], Science [21-23], Social Sciences ([24]), and Talent Management [25]), with some recent applications to Economics ([26] for the specific area of agricultural and ecological economics, [27] for a gender perspective using Portuguese institutions, and [28] for an analysis to measure academic performance in Spain).

Focusing here on evaluate the individual and collaborative merits of the Nobel laureates in Economics, we borrow several methodologies from the Complex Networks field, after assuming that the published papers are the most important merits of academics. The use of these techniques will help us to analyse the performance of the researchers, not only individually, but also as members of a community of collaborators, to estimate the importance of the co-authorship networks in obtaining better results, and to know which laureates are more, or less, collaborative.

We build the networks of collaboration denoted by co-authorship, and assign to each paper the (Eb 8703(c)3.74Jun4(b)-0.219.7162( )-10.1525(o)-0.295585(n)9.71032()-2.16436(ed(t)-2.16558( )-0.1477

The most important questions we address in the process to filter and refine have to do with the following problems. First, the same Nobel Prize laureate can use a significant number of different signatures (first name, last name, a rearrangement of them, special characters…) and, second, the same author can use different ways to specify his/her affiliation (address, centre, city…). Additionally, one author can often change affiliation, corresponding to an actual relocation from one institution to another.

To minimise errors in the identification of authors, we execute a series of tests and crosschecks (manually, in some cases). To determine that two different signatures or addresses refer to the same individual, we use the Levenshtein distance between strings [30], where distance is defined as the number of insertions or deletions needed to convert one string into another.

This unification process is performed in two steps. In the first, we run a script that takes all the different signatures of a researcher and, applying the majority rule, assigns him/her to a single research centre, a unique city, and a country. The script then analyses all authors in pairs, trying to determine if each pair corresponds to the same person. This comparison is based on the Levensthein distance of the full names used in the signatures, applying a threshold of 5% of the length of the first one. If this condition is fulfilled, then it checks whether the city of both authors is the same, and if it is, they are treated as a single entity.

The second step is performed manually. We export all authors to a CSV (Comma Separated Values) file, one per row, and we place the name, country, city, and centre, one per column. In this way, we can check, one by one, each Nobel Prize winner and his/her attributes, and see if there is a match with real information we can find on the Internet (e.g. Wikipedia).

metrics, based on the IF, while the rest can be seen in the online complementary information on http://research.kampal.com/visualization/nobel-de-economia/.

As is well known, the IF depends on the number of researchers publishing in the collaborations, and lesser or weaker collaborations outside the group, which correspond to the intuitive concept of communities. In order to gain a precise determination of these communities and to do so in an automatic way, we use walktrap [33] and leading-eigenvector algorithms [34]. The latter is used for very large networks (>10,000 nodes) in order to reduce the computing time; for the present study, only the former has been necessary.

We also define different kinds of centrality measures to quantify which are the most cohesive nodes, or those with the greatest authority [34]. In this paper, we use the betweenness, the importance of a node to connect different communities, and the Page Rank centrality, related to the number of important nodes that point to it [35].

Network Analysis Results

Nobel Laureates Network

From 1969 to 2016, 78 economists from different disciplines have been awarded the Nobel Prize. Starting from a simple geographical analysis based on the country of ascription (when a researcher has had several affiliations, we take the one with the larger number of publications), it is easy to see that the USA clearly dominates the awards, with 57 prizes, followed by the UK with 7 laureates. Other countries with Nobel laurates are Norway, Germany, France, Israel, Russia, Sweden, India, and the Netherlands.

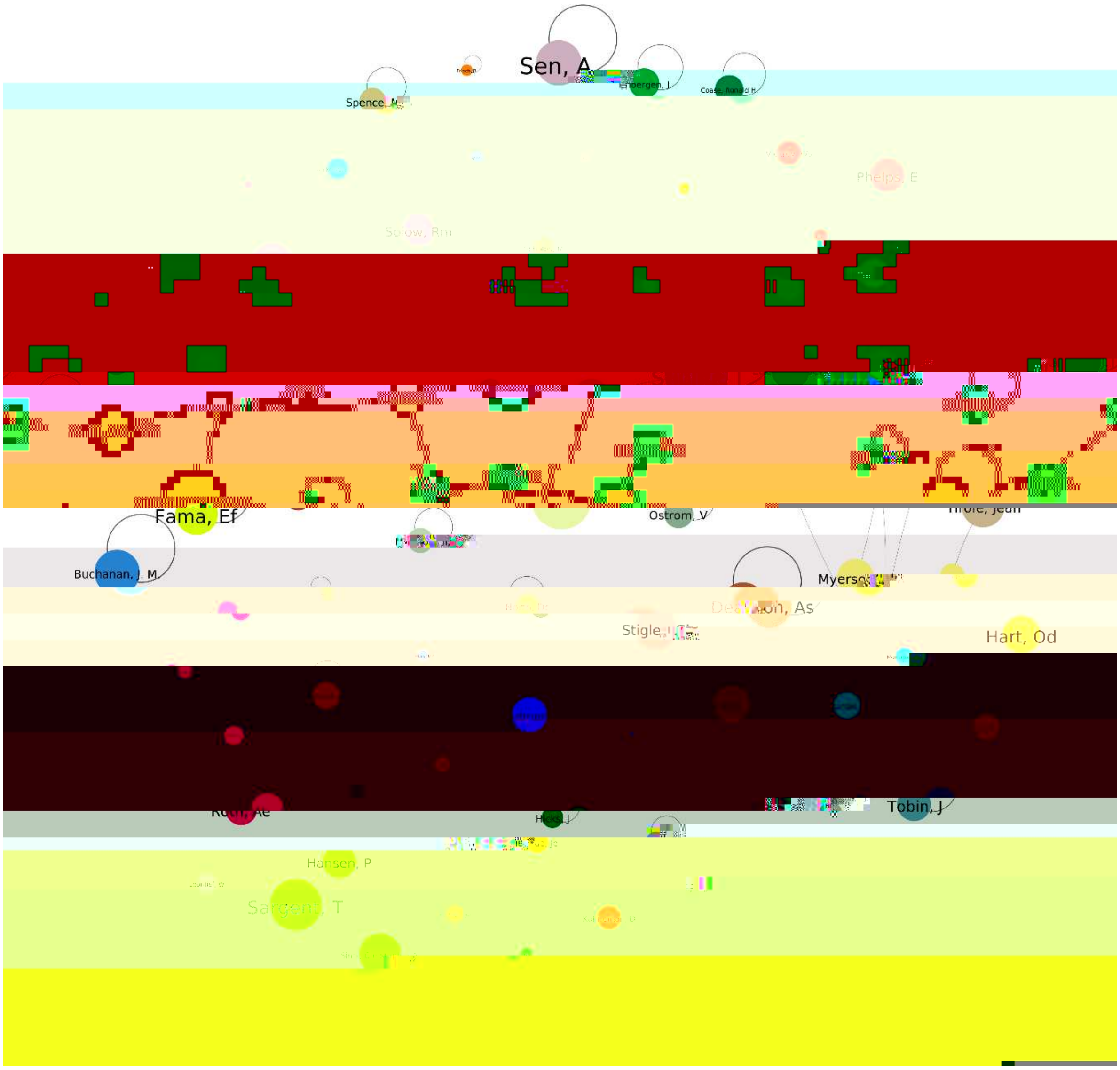
When analyzing the evolution of the production of Nobel laureates over time, defined as the sum of the NIF of the articles published by any laureate for each year, we obtain the results shown in Fig 1. It is curious to see that there is an increase ahead of the most important financial and economic crises, an indication, in some way, of an “exciting” economics field that, by anticipating the crisis, could encourage the production of Nobel laureates, with a decrease or stabilization a few years later (it should be remembered that the entire process - from the research idea up to publication - requires several years).

In Table 1, we present the top 10 researchers according to their total Normalized Impact Factor. Several of these have only recently been awarded (2011-2016), others correspond to the period 2000-2002, while others were awarded in the first years (1970-1972). Those years correspond, in fact, to the most prolific periods, according to the time evolution described above. The authors involved are experts in microeconomics, macroeconomics, or econometrics, with no clearly predominant focus.

Figure 1. Time evolution of the Nobel laureates’ Normalized Impact Factor over the years

Table 1. Top 10 Nobel laureates according to their total Normalized Impact Factor

Figure 2. Nobel laureates network (nodes colored according to the automatically-detected communities)



One can see that direct collaborations between Nobel laureates are in general rare, although there exist certain subgroups of researchers who do form connected clusters. In particular, on the central and right region of the figure, there is a rather large connected group, led by Stiglitz, and formed by 18 economists with an economic theory focus, thus including both microeconomists and macroeconomists, such as Smith, Samuelson, Lucas, Maskin, Tirole, Myerson, and Hart, among others. Other lower clusters are led by authors from the mathematical economy area, such as Roth, including Selten, Auman, and Shaply; by Sargent, with two other members, with Sims and Hansen; and, finally, by Arrow, with he being the leader of Solow and McFadden.

Following these initial analyses, we can ask whether these individual efforts have something to do with the way collaboration takes place with other researchers, and this is done in the following section.

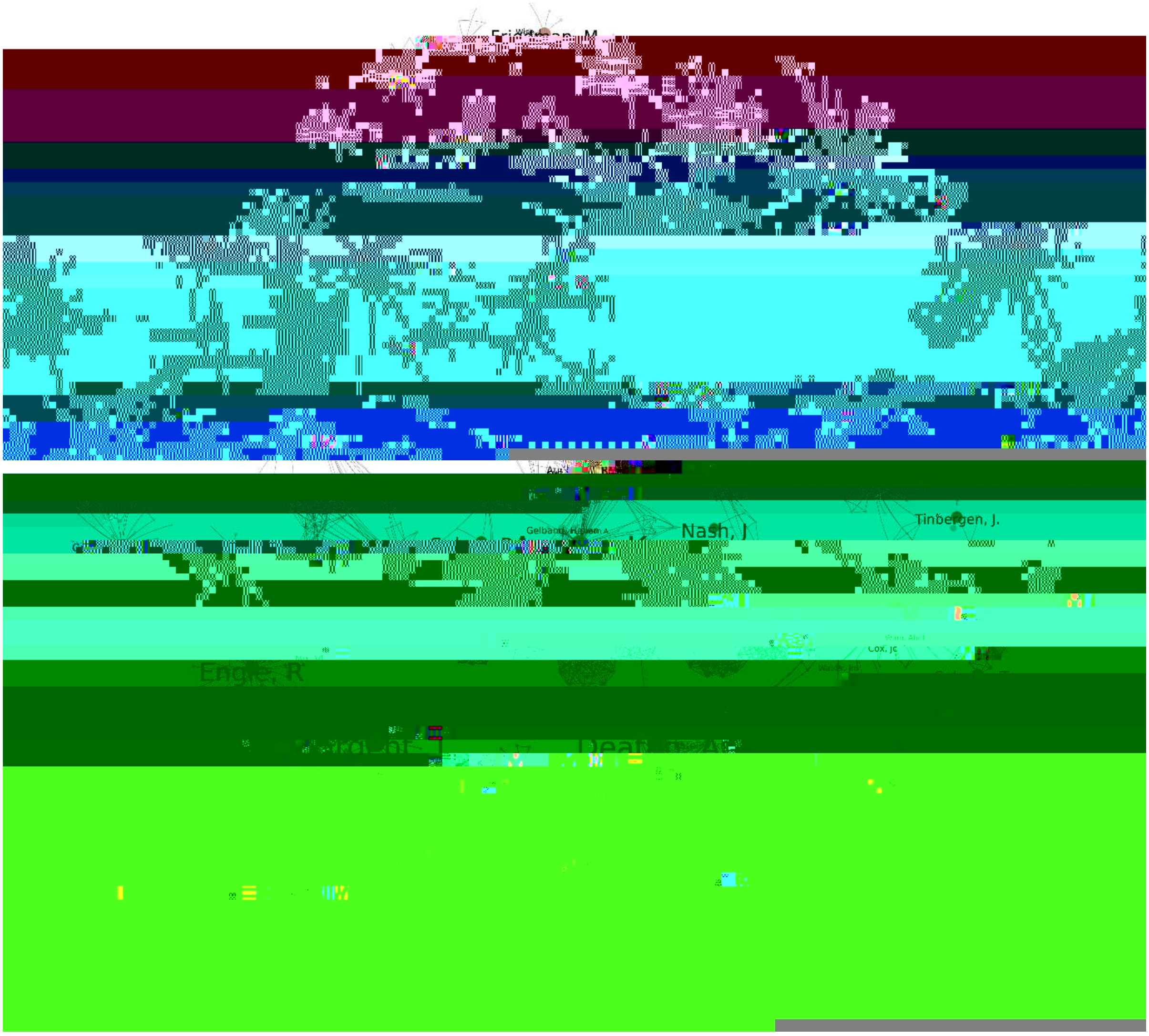
Nobel Laureates and Collaborators

We now include in the network, not only the Nobel laureates, but also their collaborators (taking into account that, for the collaborators, we consider only the work done in collaboration with Nobel laureates, as our data scope includes only papers signed by at least one prize-winner). With this, we obtain a much richer network, with a total of 1,015 researchers and a larger number of connections (see Fig 3).

The number of researchers in the large component is 715 (70% of the nodes), showing that it is a more connected network than the previous one, though the modularity is large (0.90), indicating that collaborations between the different groups is still weak.

There are certain researchers who build bridges between those groups, and this ability can be quantified through the betweenness. When we do so (see Table 3), we observe that, among the authors leading the betweenness ranking, are Arrow, Modigliani, Miller, and Tirole, laureates with a large production, with a significant number of collaborations, and with a very central position in the network. But we see

Figure 3. Network formed by the Nobel laureates (in color) and their collaborators (gray nodes)



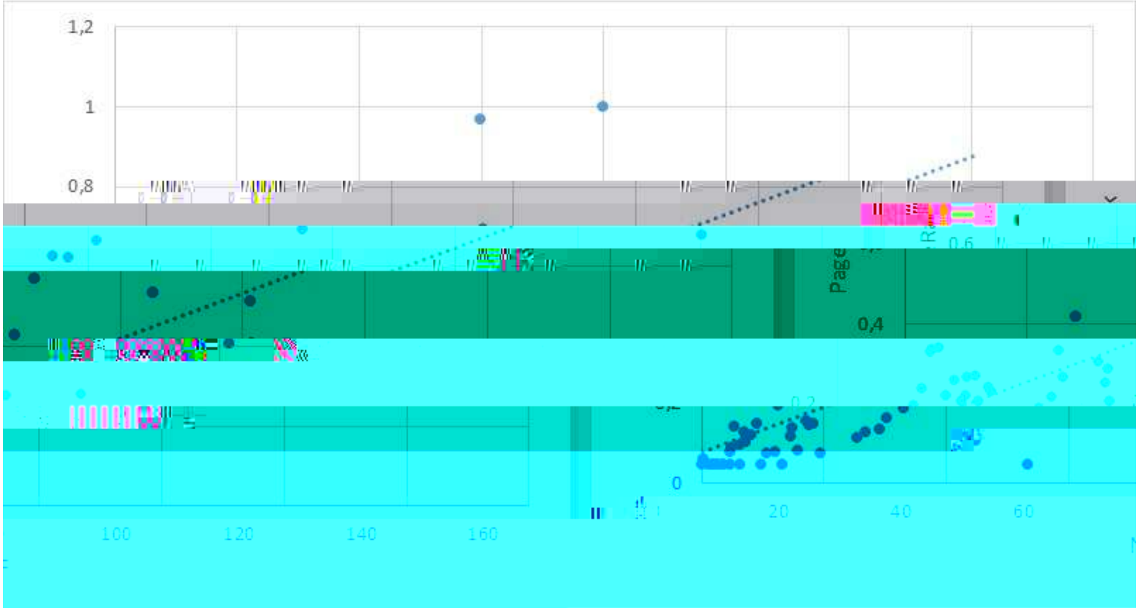
also that the top position is occupied by Grossman, a non-laureate with a smaller production in the network (remember that, for the authors who have not been laureates, only the production carried out in collaboration with Nobel winners is considered here) that, however, plays a relevant role, giving consistency to the network because he joins important parts of it. Grossman has collaborations with Stiglitz, Hart, and Shiller, among others.

Table 3. Top 10 researchers according to the betweenness in the network

On the basis of all of the above, one question arises: is there a relationship between the production of the researchers and their level of collaboration or their position in the network?

Table

Figure 4. Scatter plot of the production (NIF) and page rank of the Nobel laureates where we have added the result of a linear regression



On the other hand, the positioning algorithms and the automatic detection of communities in the network give the results presented in Fig 5. Most of the communities detected in this way are associated with one of the laureates, though some of them include more than one. When we give a name to each community according to its more productive researcher (the one with the largest NIF), we find that the 10 communities with the larger total production are those presented in Table 8. All are associated with "leaders" who have a significant individual production. However, the internal structure of those communities can be very different. In fact, we note, for example, that Deaton and Arrow are surrounded by many very productive researchers who are not laureates. The communities of Hard and Tirole, on the other hand, include several other laureates (Hart, Shiller, Miller in the first; Tirole, Maskin, Myerson, Holmstroem in the second) with a similar level of supremacy. And there are other groups that have a very hierarchical structure, with a powerful leader and a series of collaborators with a secondary role (let us remember, once again, that the production of the non-laureates is not fully considered in this study). Examples of this last case are the communities of Heckmann, Stiglitz, and Smith.

Table 8. Leaders, production and number of members of the main automatically-detected communities, according to their total NIF

Figure 5. Graphical representation of the network colored by the automatically-detected communities

We sum up the NIF of the publications of each laure

Figure 6. Color map of the NIF per year (X axis) and author (Y axis)

What we want to convey from the figure is not so much the individual features (which are not readily discernible) as the global patterns, especially with regard to the awarding year (which roughly corresponds to the upper left – lower right diagonal). It can be seen that the 1970s and the last ten years have been the most productive (more non-blue areas on the figure). But probably the pattern that is most obvious to the naked eye is that the upper part of the figure, corresponding to the early years of the awards, is rather different to the lower one (recent years).

To quantify whether this appreciation has, in fact, some statistical relevance when related to specific award years, we have calculated the evolution of the production of the laureates as a function of the time distance to the awarding year. We have estimated this production as a 3-year moving average of the NIF, starting in the year in question. This helps to smooth the curves (the output of an individual year for a small group of researchers presents rather large fluctuations), and 3 years seems to be a reasonable period, taking into account that the typical elaboration of a paper can take between 1 and 3 years from the initial idea up to publication. The results of this analysis are shown in Fig 7.

Looking at the curve corresponding to the whole set of laureates, one can see a relatively stable plateau in the years long before the award date, an increase in the years previous to it, a decrease just after the award with a recovery afterwards, followed by a new gradual decrease as the years go by - and an, at first sight surprising, peak of the research output at the end of the career. With this last exception, the results appear consistent with the different phases in the life of a researcher who, at some moment, has been awarded such an important prize.

However, when comparing the curves corresponding to the laureates up to 1997 with those from 1998, several differences can be appreciated. After the initial plateau, which is similar in both cases, the first presents a gradual decrease leading to a smaller

impact of an article is counted for each author) and can help the authors to cross over into other disciplines or fields of research.

When looking at the evolution of the research careers of the laureates from the point of view of their publications, we find significant differences between most of those awarded up to the mid-1990s and those awarded afterwards. In the first case, the career is more stable, with a gradual decrease after the award, while in the second the winners experience a sharp growth of the IF before the prize, a decrease during the years immediately after, and a new increase after that, returning to high levels of impact.

Acknowledgements

This paper was partially written while Jose Alberto Molina was Visiting Fellow at the Department of Economics of Boston College (US), to which he would like to express his thanks for the hospitality and facilities provided. Kampal Data Solutions S.L. thanks the Web of Science for permission to publish the analysis of these data on its web page (research.kampal.com). We want to thank Alfredo Ferrer for his help in some aspects of the data treatment and visualization. This paper has benefited from funding from the Spanish Ministry of Economics (FIS2015-65078-C2-2-P).

References

1. Zuckerman, H (1967) Nobel laureates in science: patters of productivity, collaboration, and authorship. American Sociological Review 32: 391-403.
2. Merton, R K (1968) The Mattew Effect in Science. Science 159: 56-63.
3. Barabási, A-L, Albert, R (1999) Emergence of scaling in random networks. Science 286: 509-512.
4. Glänzel W, Schubert A (2004) Analyzing scientific networks through co-authorship. In: Moed, H. F., Glänzel, W., Schmoch, U. (Eds), Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems, Kluwer, Dordrecht, 257–276.
5. Inzelt A, Schubert A, Schubert M (2009) Incremental citation impact due to international co-authorship in Hugarian higher education institutions. Scientometrics 78: 37-43.
6. Padial AA, Nabout JC, Wiqueira T, Bini LM, Diniz-Filho JAF (2010) Weak evidence for determinants of citation frequency in ecological articles. Scientometrics 85: 1-12.
7. Sahu SR, Panda KC (2014) Does the Multi-authorship Trend Influence the Quality of an Article? Scientometrics 98: 2161-2168.
8. Schubert A (2014) Sentences to remember from the first 100 volumes of the Journal Scientometrics. Scientometrics 100: 1-13.
9. Ductor L (2016) Does co-authorship lead to higher academic productivity? Oxford Bulletin of Economics and Statistics 77: 385-407.
10. Sutter M, Kocher M (2004) Patterns of co-authorship among Economics Departments in the USA. Applied Economics 36: 327-333.

11. Goyal S, Van Der Leij MJ, Moraga-González JL (2006) Economics: an emerging small world. Journal of Political Economy 114: 403-412.

12. Rath K, Wohlrabe K (2016) Recent trends in co-authorship in Economics: evidence from RePEc. Applied Economics Letters 23: 897-902.

13. Wagner CS, Horlings E, Whetsell TA, Mattsson P, Nordqvist K. Do Nobel laureates create prize-winning networks? An analysis of collaborative research in Psysiology or Medicine. PLoS ONE 10(7): e0134164.

14. Parreira M, Machado K, Logares R, Diniz-Filho J, Nabout JC (2017) The roles of geographic distance and socioeconomic factors on international collaboration among ecologists. Scientometrics 113: 1539-1550.

15. Colavizza G (2017) The structural role of the core literature in history. Scientometrics 113: 1787-1809.

16. Tang M, Cheng Y, Chen K (2017) A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses. Scientometrics 113: 985-1008.

17. Robert C, Arreto, C, Azerad J, Gaudy J (2004) Bibliometric overview of the utilization of artificial neural networks in medicine and biology. Scientometrics 59:117-130.

18. Rainho O, Cointet J, Cambrosio A (2017) Oncology research in late twentieth century and turn of the century Portugal: a scientometric approach to its institutional and semantic dimensions. Scientometrics 113: 867-888.

19. Bergé L, Scherngell T, Wanzenböck I (2017) Brinding centrality as an indicator to measure the “brinding role” of actors in networks: An application to the European Nanotechnology co-publication network. Journal of Informetrics 11: 1031-1042.

20.

28. Molina JA, Ferrer A, Iñiguez D, Rivero A, Ruiz G, Tarancón A (2018) Network analysis to measure academic performance in Economics. DOI: 10.1007/s00181-1546-0.

29. Alvarez R, Cauhé E, Clemente-gallardo J, Ferrer A, Iñiguez D, Mellado X, Rivero A, Ruiz G, Sanz F, Serrano E, Tarancón A, Vergara Y (2015) Analysis of academic productivity based on Complex Networks. Scientometrics 104: 651-672.

30. Levenshtein I (1996) Binary codes capable of correcting deletions, insertions and reversals, Cybernetics and Control Theory 10: 7076710.

31. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex Networks: Structure and Dynamics. Physics Reports 424: 175-308.

32. Fruchterman TMJ, Reingold EM (1991) Graph Drawing by Force-directed. Software: Practice and Experience, 21(11): 1129.

33.