

Molina, José Alberto; Iñíguez, David; Ruiz, Gonzalo; Tarancón, Alfonso

Working Paper

Networks in Population Economics: production and collaborations

GLO Discussion Paper, No. 1051

Provided in Cooperation with:

Global Labor Organization (GLO)

Suggested Citation: Molina, José Alberto; Iñíguez, David; Ruiz, Gonzalo; Tarancón, Alfonso (2022) : Networks in Population Economics: production and collaborations, GLO Discussion Paper, No. 1051, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/249774>

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.

Networks in Population Economics: production and collaborations

José Alberto Molina^{1,4,5*}, David Iñíguez^{2,3,6}, Gonzalo Ruiz⁶, Alfonso Tarancón^{2,6}

1. Department of Economics, University of Zaragoza, Spain
2. Department of Theoretical Physics, University of Zaragoza, Spain
3. ARAID Foundation, Diputación General de Aragón, Spain
4. Global Labor Organization (GLO), The Netherlands
5. Institute on Employment, Digital Society and Sustainability (IEDIS), University of Zaragoza, Spain
6. Institute for Biocomputing and Physics of Complex Networks (BIFI), University of Zaragoza, Spain

Abstract: Population Economics (PopEc) covers a number of topics in Economics, as well as in Demography, Labor Studies, Sociology, etc. For example, the economic determinants of population change and demographic behavior covers topics such as household formation, marriage and divorce, fertility, gender, child bearing, schooling, access to labor markets, migration, well-being, and ageing and mortality, among others. In this paper, we analyze the production and networks of a total of 6,472 authors who have published 5,070 papers in Population Journals (indexed in SSCI or WOS) between January 1969 and January 2021 (Journal of Population Economics, Journal of Human Resources, Feminist Economics, the Review of Economics of the Household, the Journal of Demographic Economics, Demography, Population and Development Review, and the European Journal of Population). Using the Impact Factor (IF) of these journals corresponding to the year of publication, our results first identify the academic leaders among those authors, as well as other results in terms of communities. Results reveal that the largest community is led by the sociologist Trude Lappegard, with the community led by the economist Hans-Peter Kohler as a close second. The latter community includes the most prolific author, the economist Samuel H. Preston. Additionally, we note that collaborations among these authors are very rare, with only their neighborhoods collaborating.

Keywords: Population Economists; Impact factor; Leaders and co-authorship; Research production; Complex networks

JEL codes: A11, C45

* Corresponding author: JA Molina. e-mail: jamolina@unizar.es

1. Introduction

Population Economics (PopEc) covers a number of topics in Economics, as well as in Demography, Labor Studies, Sociology, etc. For example, the economic determinants of population change and demographic behavior covers topics such as household formation, marriage and divorce, fertility, gender, child bearing, schooling, access to labor markets, migration, well-being, and ageing and mortality, among others. Articles in Population Economics are usually published in an extensive number of journals, with several journals being indexed as top SSCI journals. Figure 1 shows the number of articles published in a select group of SSCI journals, which can be included in the category of Population Economics: Journal of Human Resources, Journal of Population Economics, Feminist Economics, the Review of Economics of the Household, the Journal of Demographic Economics, Demography, Population and Development Review, and the European Journal of Population. Despite the increasing trend, particularly since 2008, of papers published in these journals (with a main explanatory factor being the establishment of the Review of Economics of the Household), no analyses of collaborations and production of participating Economists has been, up to now, undertaken.

(Figure 1 about here)

This paper partially bridges this gap, providing new insight into collaborations and outcomes of Population Economists by developing network analysis of those authors who have published in the selected PopEc journals, in order to illustrate their interrelations and partnerships. We identify five journals from SSCI-Economics (Journal of Human Resources, Journal of Population Economics, Feminist Economics, Review of Economics of the Household, the Journal of Demographic Economics) and three journals from SSCI-Demography (Demography, Population and Development Review, and European Journal of Population). We evaluate co-authorship structures and identify natural communities, hubs, and leaders in terms of centrality and influence (Newman 2010). Additionally, we perform a specific analysis of those who have published in the Journal of Population Economics. To perform the analyses, as well as for the graphical representations of the networks, we use the application Kampal Research (Alvarez et al., 2015).

The analysis of collaborative academic networks has emerged from Physics and Computational sciences (see, e.g., Price, 1965 and Cobo et al., 2011), looking for answers to questions such as “who are the researchers with a greater central role in the network, or who encourage more cohesion among the members of the network”. Network theory is a very useful framework in which to characterize the functioning of complex systems (Newman 2010; Estrada 2011; Latora et al. 2016), providing a versatile representation of the interaction scheme between the different elements, where powerful tools from statistical physics (Albert & Barabasi 2002) and dynamical systems (Barrat & Vespignani 2008) can be used to study their functional and structural properties.

In this paper, we apply complex network techniques to the set of authors who have published in Population Economics. The economics literature shows the importance of the evaluation of networks and productivity, with the general bibliography showing that collaboration in writing papers confers several advantages on co-authors, in terms of more citations (see Molina et al. 2020 and 2021 for a review of this literature). In particular, complex networks have been extensively used in bibliometric analysis, with Price (1965) being a seminal paper, and Kumar (2015) being a review of the literature associated with co-authorship networks. The use of these techniques will help us to analyze the performance of the researchers, not only individually, but also as members of a community of collaborators, to determine their importance, not only from a purely statistical point of view, but also from the role they play in the topology and connectivity of the network.

To construct the network, we define individual researchers as nodes, and relations between them, generated by common articles, as links. Network analyses then allow us to discern global properties and represent the results via maps that encode the relevant information in graphical terms, as well as numerical parameters that display the topological properties of the network. One important aspect when studying collaboration networks is the weight given to the links, as explained in Perianes-Rodriguez et al. (2016). In our case, as described below, we base this weight on the JCR Impact Factor (IF) of the articles published in common. To generate informative and comprehensible interaction analyses, a considerable amount of work is needed to obtain and ‘clean’ the data; that is to say, the identification of authors is not automatic, and there are bugs due to varying details in signatures and affiliations. Another important aspect, which we examine closely,

is the complementarity that can be obtained from a joint analysis of the production and the centrality of researchers in the network.

We complete the study with an analysis of the main topics of interest present in the journals, not based on previously defined categories but on automatically identified “keywords” (words or text chains with a semantic content) extracted from the title and abstract of the publications. Many text analysis tools are based on classical TF-IDF algorithms (Luhn 1957, Spark Jones 1972, Robertson 2004) that are complemented with certain vector-mapping techniques, to calculate semantic textual similarity (Le & Mikolov 2014). We use a combination of both approaches in order to find a precise set of words or text chains that can be identified as the topics of work of the different authors.

The organization of the paper is as follows. Section 2 describes the data scope and processing. Sections 3, 4, and 5 describe the analysis and results for the whole set of PopEc journals, while Section 6 is focused on the Journal of Population Economics (JOPE). In particular, Section 3 presents the network approach and the corresponding results for the author network, Section 4 shows a country analysis, and Section 5 is centered on the collaboration communities and main topics of research. We close the paper with our main conclusions.

2. Data

Our data have been extracted from the Web of ScienceTM (WoS) in a group of JCR-ECON journals, with especial emphasis on PopEc, encompassing a total of 5,070 papers published between January 1969 and January 2021, according to the following distribution: Journal of Human Resources (2,178 papers from 1966), Journal of Population Economics (799 papers from 1992), Feminist Economics (987 papers from 1998), the Review of Economics of the Household (487 papers from 2008), the Journal of Demographic Economics (99 papers from 2015), Demography (3,124 papers from 1864), Population and Development Review (3,464 papers from 1975), and European Journal of Population (165 papers from 1985).

There are difficulties in unequivocally identifying authors, mainly due to incoherencies in signatures or affiliations, but we have applied a series of processes to do so with a reasonable level

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

To clearly identify the authors, we perform a process of unification, an evolution of the method described in Molina et al. (2020, 2021). In this case, we develop a general algorithm to automatically implement the human tasks of the previous work. We process the following maximal set of data for each initial (not yet unique) author: Researcher_Id, email, name, surname, name initials, city, country, affiliations, keywords. However, this information is incomplete; in general, for each author we only know a subset of this maximal list of variables. In particular, the Researcher_Id field, which in principle would directly identify a unique author, is available only in 10% of the cases. Furthermore, there are several researchers who have more than one Researcher_Id.

To verify whether two authors in this initial list are the same, we compute a distance between them, a weighted sum of distances between each of the different fields (Researcher_Id, email, etc.). To compute these distances, we use different algorithms. For strings, we use the Levenshtein distance (Levenshtein 1966). For cities and countries, first we compare them with an international corpus, to obtain a unique text. For names and surnames, we extract the name initials (if they are not already in this format); next we compare initials, names, and surnames and compute a distance, which is zero if both are the same. Finally, we calculate the total distance between the two candidate authors through a linear combination of these elemental distances, where the coefficients differ depending on the capacity of the corresponding variable to discriminate between different authors (for example, the email coefficient is bigger than the institution one). If this total distance is smaller than a predefined limit, we assume that the two authors are the same and unify them, adding cities, countries, affiliations, and keywords. In this way, we obtain 6,472 unique authors (from a total of 10,083 initially present) who have published 5,070 papers between January 1969 and January 2021.

3. Authors' Networks

We construct and analyze the network of researchers by applying methodologies from the Complex Networks discipline (see Boccaletti et al., 2006, and, specifically, Alvarez et al., 2015, for a more complete description of the specific tools and procedures we use). To build a network, we must define nodes and links. Here, the nodes will be the researchers under study (the authors who have published in the PopEc selected journals), while the links between any two nodes will be defined by the collaborations (common publications) and the weights of these links will be related to the strength of these relationships.

A range of metrics are commonly used to measure the relevance of publications and, in consequence, of the authors; the number of articles, JCR impact factor (IF), quartiles, number of excellence papers, etc. Here, since the set of PopEc journals belongs to similar JCR categories, we will use the IF to weigh the importance of a node (author) and of a link (connection between two authors). Thus, the weight associated with a link between two authors - the common publication of an article in a certain journal - is its IF divided by its total number of co-authors. We include a self-link with the same weight, so, when we sum all the links generated by a particular paper for a certain author, we find that the weight associated with a given paper is in fact its IF. The total weight of the link between two authors is the sum of the weights associated with every common paper.

To graphically represent the network, we consider the system to be mechanical, with forces making the system evolve, in a similar way to a system of particles. Using force-directed algorithms (Fruchterman et al. 1991), and a Monte Carlo process to separate overlapping researchers, we obtain graphs in which nodes with a stronger relationship are closer. This provides a geometrical vision of the network which is useful to visually identify groups of researchers with stronger internal collaborations, and lesser or weaker collaborations outside the group, which correspond to the intuitive concept of communities. For the automatic detection of these communities, we use walktrap (Pons and Latapy 2006) and leading-eigenvector algorithms (Newman, 2006). 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 can define different centrality indicators in order to measure the relevance of the nodes from a topological point of view, which are the most cohesive, or those with the greatest authority (Newman, 2006). In our study, we have used the between-ness centrality (Freeman, 1977; Newman, 2005), which indicates the importance of a node to bridge different parts of the network, and the page rank centrality (Page et al., 1998), related to the number of important nodes that point to it and that provides an estimate of the importance or authority of the node within its neighborhood.

The following Tables show the most important researchers in the PopEc network, according to a range of indicators. Table 1 shows authors ordered by total production (IF), with the Princeton economist, demographer and sociologist Samuel H. Preston occupying the first position. John Bongaarts is a Dutch-American demographer, and the Ph. D. in Economics Ronald D. Lee are second and third in this ranking. Five of the top 10 are economists, with the rest being demographers and sociologists. Table 2 similarly presents authors ordered by their between-ness or centrality, which indicates the capacity to build bridges between different groups. The psychologist Ariel Kalil is in the first position, with the sociologist Francis Goldscheider appearing second. Table 3 similarly shows authors by the indicator page rank, which constitutes a more local definition of centrality than between-ness, indicating the importance of the node in its neighborhood, with the economists Samuel H. Preston and Jere R. Behrman appearing in the first positions.

(Tables 1, 2, 3 about here)

The network formed by all the PopEc authors appears in Figure 2 (the size of the nodes is given by the IF of the researcher, while their color corresponds to the automatically detected communities). The largest connected cluster (giant cluster) consists of 1,941 researchers (approximately 30% of the total), showing that, in the subset of journals considered, a large collaborative structure is present, despite that many isolated communities also appear. In this Figure, we identify the academics who appear in the first position in the previous three tables: the economist Samuel H. Preston and the psychologist Ariel Kalil.

(Figure 2 about here)

Additionally, we explore whether there is a relationship between the production in terms of IF of the researchers and their level of collaboration, or their position in the network. The Pearson correlations between the production of the authors and their between-ness, page rank, and number of collaborators are shown in Table 4. These three values of correlations are relatively large, although with significant differences. In fact, the correlation of production with between-ness is weaker, because the ability to build bridges between different regions of the network does not necessarily guarantee a much better performance. However, the two other magnitudes, more directly related to the local collaborative activity, present a strong correlation with production, both the number of collaborators and, especially, the page rank. In Fig 3, we show a scatter plot of the production and page rank of the researchers where, despite the element of dispersion, the strong correlations can be appreciated.

(Table 4 about here) (Figure 3 about here)

4. Country Networks

For global activity by countries, as well as collaborations among researchers from different countries, we construct the Country Network, by grouping all researchers with the same country affiliation. The country production is the sum of the production of all members. We obtain a node for each country, and then create a link between two countries with weight proportional to the number of collaborations between countries (Figure 4). The US, the UK, and Germany are the most productive countries in PopEc. In the same Figure, we compute the Countries Communities, that is to say, the set of countries that collaborates internally more than the others, with the countries in the same community having the same color. We identify three clear groups: the close collaborations between US, Australia, and China, between Germany and the Nordic countries, and, finally, the collaborations of the UK with the Mediterranean countries.

(Figure 4 about here)

Table 5 shows the production for each country. Note that a paper authored by, e.g., two authors from different countries is added to both of them. The USA production is the largest one (total

impact 9,250), very much larger than the second (United Kingdom, 861) and the third one (Germany, 741). Centralities are also important because they give us information about the role played in uniting the global network. The most cohesive countries are USA and UK (as expected from their high production) and Canada, but the latter's relevance originates in its collaborations with a large number of other countries, rather than being based in a high level of production.

(Table 5 about here)

5. Communities and main topics of research

In the previous section, we focused on quantifying the activity of individual researchers in terms of their production (IF) and of their position in the network (centralities). But the questions remain, how do they collaborate with each other and what do they work on?

The first step is to extract from the title and abstract the main keywords or topics of interest (words or text chains with a semantic content) of each publication, and then associate these with the authors. To that end, we use a TF-IDF algorithm to extract a first list of keywords that we refine by mapping them to a semantic space, using a BERT neural network that works as a sentence encoder (Reimers and Gurevych 2019). In this way, we find a precise set of words or text chain that can be identified as the topics of work of the different authors.

Table 6 shows the most frequent topics of the papers, appearing in the biggest countries (USA, India, China) and other geographical sets, such as sub-Saharan Africa. If we remove them from the list, the topics with the largest presence in the set of journals selected are fertility, children, labor supply, and marriage and mortality, in this way identifying the most important topics analyzed in our sample journals. When we do the same exercise for other countries of assignment, we appreciate clear differences among their main topics of interest. In the US, and due to the large quantity of publications, the topics are quite diverse, with some pre-eminence of demography. In the case of UK, the focus seems to be more influenced by geographical and cultural factors than by some specific topic, as many English-speaking countries are present in the list. In Germany, we

identify the family- and demography-related topics, whereas Canada is dominated by health-related topics. Additionally, results show that Italy highlights education and life quality topics, the family economy topics dominate in Spain, and the main topics in France are related to immigration.

(Table 6 about here)

To examine the topics at the communities level, we focus on those detected by the algorithms cited in the previous section. When we name each community according to its most productive researcher (the one with the largest IF), we find that the 10 communities with the largest total outcome are those presented in Table 7. All are associated with “leaders” who have significant individual production. However, the internal structure of those communities can differ significantly. The largest is led by the sociologist Trude Lappegard, the community with the largest production is led by the economist Samuel H. Preston, and the one associated with the most central author, the psychologist Ariel Kalil, is led by the economist Greg J. Duncan.

(Table 7 about here)

Lappegard’s is a large community (152 members), with many authors having a similar orientation (see Fig. 5). The main topics of interest are related to specific European countries, with some focus on family matters (cohabitation, birth order, divorce...). Preston’s is a much more compact community, with 81 members and a more hierarchical structure, with Preston standing out among the rest in terms of outcome (see Fig. 6). The topics in this community are more related to the US and some developing countries, in fields such as intimate partner violence, attitudes, and life expectancy. Duncan’s community seems to be a middle ground, with 72 members but a less hierarchical structure, with Duncan and Kalil having a slightly bigger influence than the rest (see Fig. 7). In this case, the focus of their work is on child-related subjects.

(Figure 5, 6 and 7 about here)

We have also analyzed the network formed by the top 10 authors (according to their IF) and all their coauthors (first neighbors). The result can be seen in Figure 8, where color indicates the different groups detected by the community detection algorithm. We can see that the collaborations

among these authors are very small. Only Hans-Peter Kohler and Jere R. Behrman have some direct collaboration, while the rest of the co-authorships are among the neighborhoods. The country structure of this network is very striking in that almost all authors in this excellence network are from the US, as can be seen in Figure 9, with the color indicating the author's country (violet corresponds to the US).

(Figures 8 and 9 about here)

6. JOPE Network

In the particular case of the Journal of Population Economics (JOPE), after the process of author unification explained above, the initial number of 1,520 researchers is reduced to 1,246 unique authors who have published 799 papers, between January 1992 and January 2021. Our analyses here are similar to those described in the previous sections for the set of PopEc journals.

The authors that have accumulated the largest IF with their publications in JOPE are listed in Table 8, with all being Economists. We also show the centralities of those researchers, though in this network the magnitudes are not so important, since there is no significant connected component. The network restricted to JOPE is represented in Figure 10. We can see a number of small communities, but no large, connected structure. In this Figure we have identified Pierre Pestieau as the academic who is in the first position according to the outcome (IF).

(Table 8 about here) (Figure 10 about here)

In JOPE the production is, again, led by the United States, followed by Germany and the UK (Table 9). The main topics of interest are shown in Table 10, with Fertility, Children, Growth, Labor Supply, and Education appearing in the first positions, apart from the countries. Table 11 presents the communities of the JOPE network, led by the economists Pierre Pestieau and Stephen J. Trejo. Fig 11 shows the network of the Pierre Pestieau community and its main topics of interest (Dependent parent, Altruism, and Family norms, among others).

(Tables 9, 10, 11 about here) (Figure 11 about here)

7. Conclusions

This paper provides new insight into the collaborations and production of Population Economists by developing a network analysis of authors who have published in seven SSCI journals (Journal of Human Resources, Journal of Population Economics, Feminist Economics, Review of Economics of the Household, Journal of Demographic Economics, Demography, Population and Development Review, and European Journal of Population). We identify 6,472 authors who have published 5,070 papers in the journals between January 1969 and January 2021. Additionally, we perform a specific analysis of authors who have published in the Journal of Population Economics.

Using the Impact Factor (IF) of these journals corresponding to the year of publication, our results first indicate the authors ordered by total production (IF), with the economist Samuel H. Preston, occupying the first position. John Bongaarts, a Dutch-American demographer, and the economist Ronald D. Lee follow. Five of the top 10 are economists, with the rest being demographers and sociologists. With respect to the authors ordered by their between-ness or centrality, the psychologist Ariel Kalil is in the first position. The network formed by all the PopEc authors shows that the largest connected cluster (a giant cluster) consists of 1,941 researchers (approximately 30% of the total). Additionally, the Pearson correlations between the production of the authors and their between-ness, page rank, and number of collaborators are relatively large, though with significant differences. In fact, the correlation of production with between-ness is weaker. However, the two other magnitudes, more directly related to the local collaborative activity, present a strong correlation with production, both the number of collaborators and, especially, the page rank.

The US, the UK, and Germany are the most productive countries in PopEcs, with three communities that show the close collaborations between the US, Australia, and China, between Germany and the Nordic countries and, finally, the collaborations of the UK with the Mediterranean countries. The most frequent topics, after removing the names of countries from the list, are fertility, children, labor supply, marriage, and mortality.

The 10 communities with the largest total outcome are associated with “leaders” who have significant individual production. The largest is led by the sociologist Trude Lappegard, followed by the community led by the author with the largest production, the economist Samuel H. Preston. Lappegard’s is a large community (152 members), and the main topics of interest are related to specific European countries, with some focus on family matters (cohabitation, birth order, divorce...). Preston’s is a much more compact community, with 81 members and a more hierarchical structure, with Preston standing out among the rest in terms of outcome. The topics in this community are more related to the US and certain developing countries, in fields such as intimate partner violence, attitudes, and life expectancy.

Finally, we have analyzed network collaborations for authors who have published in the Journal of Population Economics, with 1,246 authors having published 799 papers between January 1992 and January 2021. The authors who have accumulated the largest IF are all Economists. Pierre Pestieau is the academic in the first position according to the outcome (IF). In JOPE, the production is, again, led by the US, followed by Germany and the UK, with the main topics of interest being fertility, children, growth, labor supply, and education.

Acknowledgements We thank Kampal Data Solutions S.L. as well as the Web of Science for permission to publish the analysis of these data on the web page research.kampal.com. We want to thank Alfredo Ferrer and J. Ignacio Giménez-Nadal for their help in some aspects of the data treatment and visualization, as well as Alejandro Rivero for his work in the keyword detection algorithms.

Funding This paper has benefited from funding from the Spanish Ministry of Science, Innovation and Universities (PGC2018-094684-B-C22E30_17R) and the Aragon Government (Research Groups E30_20R and S32_20R).

Declarations

Conflict of interest The authors declare no competing interests.

References

- Albert R, Barabási AL (2002) Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74: 47.
- Alvarez R, Cahué E, Clemente-Gallardo J, Ferrer A, Iñíguez 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.
- Barrat A, Vespignani A (2008) *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge.
- Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex Networks: Structure and Dynamics. *Physics Reports* 424: 175-308.
- Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F (2011) Science mapping software tools: review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology* 62 (7): 1382-1402.
- Estrada E (2011) *The Structure of Complex Networks: Theory and Applications*. Oxford University Press, Oxford.
- Freeman LC (1977) A set of measures of centrality based on betweenness. *Sociometry* 40(1): 35-41.
- Fruchterman TMJ, Reingold EM (1991) Graph Drawing by Force-directed. *Software: Practice and Experience*, 21(11): 1129.
- Kumar S (2015) Co-authorship networks: a review of the literature. *Aslib Journal of Information Management* 67(1): 55-73.
- Latora V, Nicosia V, Russo G (2017) *Complex Networks: Principles, Methods and Applications*. Cambridge University Press, Cambridge.
- Le Q, Mikolov T (2014) Distributed representations of sentences and documents. *International conference on machine learning*. PMLR 32(2): 1188-1196.

Levenshtein I (1996) Binary codes capable of correcting deletions, insertions and reversals, *Cybernetics and Control Theory* 10: 707-710.

Luhn HP (1957) A Statistical Approach to Mechanized Encoding and Searching of Literary Information. *IBM Journal of Research and Development* 1 (4): 309–317.

Molina JA, Ferrer A, Iñiguez D, Rivero A, Ruiz G, Tarancón A (2020) Network analysis to measure academic performance in Economics. *Empirical Economics* 58: 995-1018.

Molina JA, Iñiguez D, Ruiz G, Tarancón A (2021) Leaders among the leaders in Economics: A network analysis of the Nobel Prize laureates. *Applied Economics Letters* 28(7): 584-589.

Newman MEJ (2005) A measure of betweenness centrality based on random walks. *Social Networks* 27(1): 39-54.

Newman MEJ (2006) Finding community structure in networks using the eigenvectors of matrices. *Physical Review E* 74(3): 036104.

Newman MEJ (2010) *Networks: An introduction*. Oxford University Press, Oxford.

Page L, Brin S, Motwani R, Winograd T (1999) The PageRank citation ranking: Bringing order to the web. *Stanford InfoLab*.

Perianes-Rodriguez A, Ludo Waltman L, van Eck NJ (2016) Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics* 10: 1178-1195.

Pons P, Latapy M (2006) Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications* 10 (2): 191-218.

Price, DJ de S (1965) Networks of scientific papers. *Science* 149 (3683): 510-515.

Reimers N, Gurevych I (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Robertson S (2004) Understanding inverse document frequency: on theoretical arguments for IDF.
Journal of Documentation 60(5): 503-520.

Sparck Jones K (1972) A statistical interpretation of term specificity and its application in retrieval.
Journal of Documentation 28(1): 11-21.

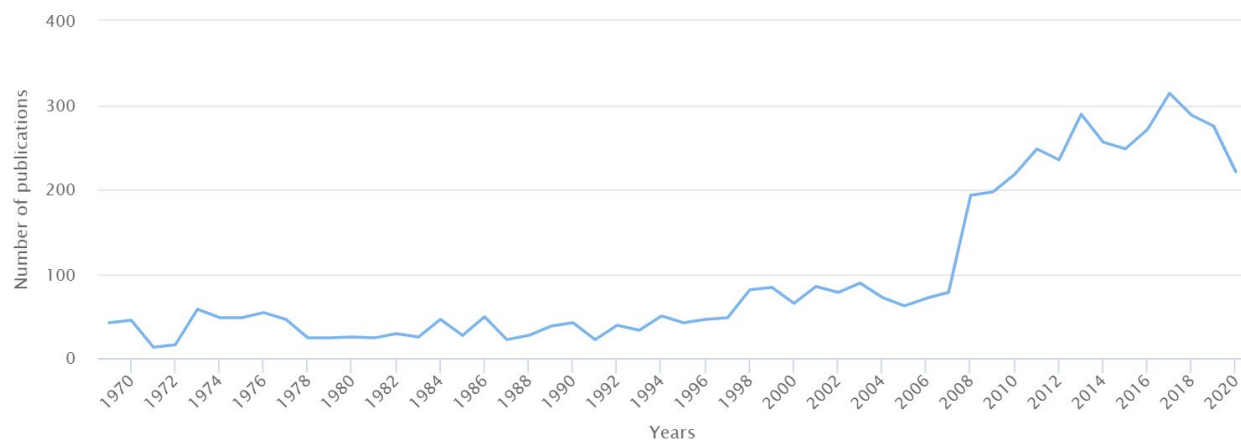


Fig. 1 Number of publications on Population Economics (annual)

Table 1 Authors ranked by Total Outcome

	Total outcome	Between-ness	Page Rank	Contributors
Samuel H. Preston (Economist)	41.94	0.60	1.00	22
John Bongaarts (Demographer)	32.04	0.28	0.64	10
Ronald D. Lee (Economist)	27.34	0.70	0.73	15
Mikko Myrskylä (Demographer)	26.13	0.31	0.55	12
S. Philip Morgan (Sociologist)	26.03	0.58	0.74	15
Alberto Palloni (Sociologist)	25.39	0.11	0.75	14
Hans-Peter Kohler (Economist)	22.89	0.77	0.73	26
Kathryn M. Yount (Demographer)	22.89	0.742	0.772	24
Jere R. Behrman (Economist)	22.32	0.90	0.79	29
Catalina Amuedo-Dorantes (Economist)	21.64	0.36	0.58	11

Table 2 Authors ranked by Between-ness

	Between-ness	Total outcome	Page Rank	Contributors
Ariel Kalil (Psychologist)	1.00	16.98	0.55	15
Fk Goldscheider (Demographer)	0.95	14.46	0.45	11
Natascha V. Pikauskas (Social Work))	0.94	10.52	0.29	6
Trude Lappegard (Sociologist)	0.93	17.43	0.50	11
Jere R. Behrman (Economist)	0.90	22.32	0.79	29
Elizabeth Thomson (Sociologist)	0.88	13.18	0.40	11
Arland Thornton (Sociologist)	0.82	15.39	0.45	18
Marcia J. Carlson (Sociologist)	0.79	7.58	0.24	8
Greg J. Duncan (Economist)	0.78	19.39	0.59	12
Hans-Peter Kohler (Economist)	0.77	22.89	0.73	26

Table 3 Authors ranked by Page Rank

	Page Rank	Total outcome	Between-ness	Contributors
Samuel H. Preston (Economist)	1.00	41.94	0.60	22
Jere R. Behrman (Economist)	0.79	22.32	0.90	29
Alberto Palloni (Sociologist)	0.75	25.39	0.11	14
S. Philip Morgan (Sociologist)	0.74	26.03	0.58	15
Hans-Peter Kohler (Economist)	0.73	22.89	0.77	26
Ronald D. Lee (Economist)	0.73	27.34	0.70	15
Kathryn M. Yount (Demographer)	0.72	22.88	0.42	24
Jeroen van Bavel (Humanities)	0.69	21.62	0.32	19
Julie Davanzo (Economist)	0.68	17.82	0.24	14
Wolfgang Lutz (Demographer)	9.66	20.69	0.14	15

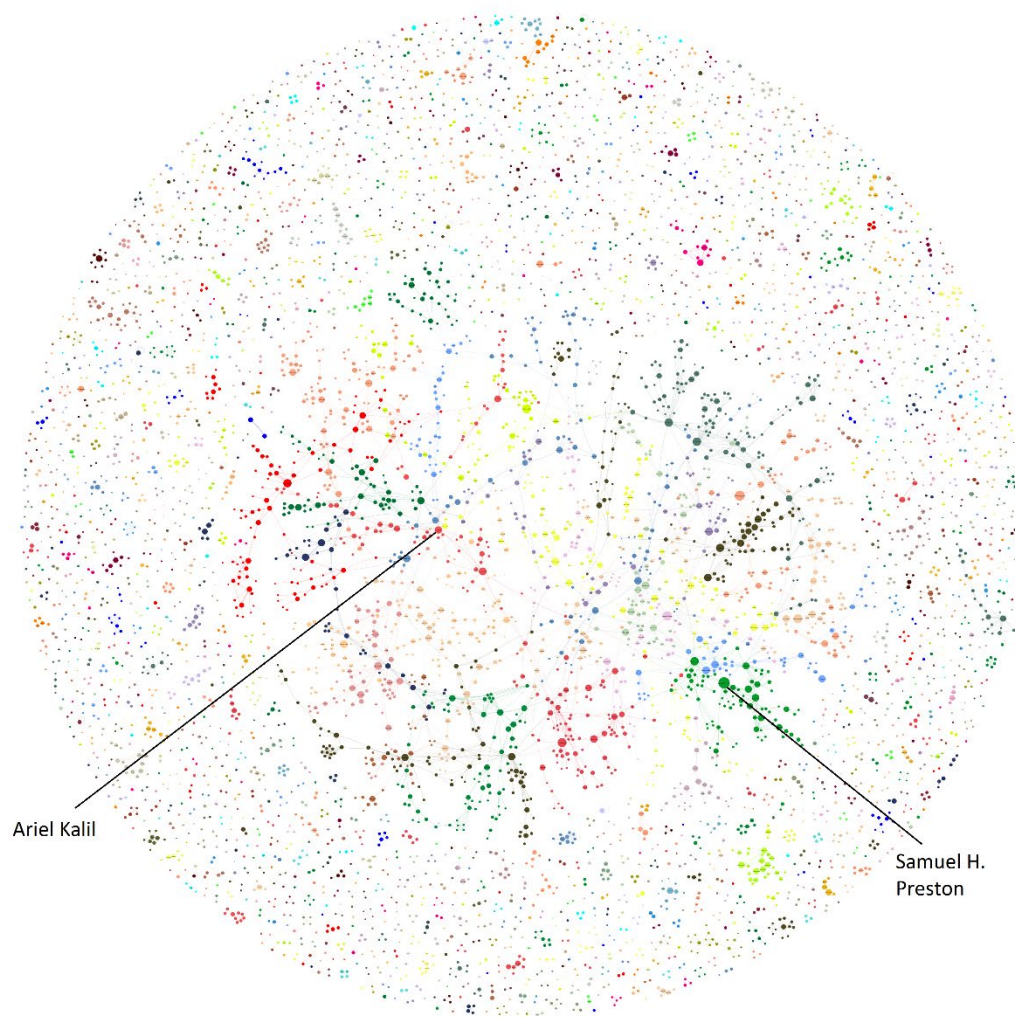


Fig. 2 Author networks

Table 4. Pearson correlation between the production (IF) of the researchers and their between-ness, page rank, and number of collaborators

	Pearson Correlation
Production/Between-ness	0.595
Production/Page Rank	0.784
Production/Collaborators	0.688

Table 5 Country production (IF)

	Total outcome	Centrality	Relevance	Contributors
United States	9250.17	1	1	63
United Kingdom	861.36	0.17099	0.24295	40
Germany	741.19	0.12558	0.22945	31
Netherlands	439.55	0.05088	0.14459	25
Canada	430.76	0.1726	0.18476	24
Italy	386.33	0	0.09609	18
Australia	377.85	0.02576	0.13544	23
Spain	367.02	0	0.08456	17
Sweden	321.39	0.02576	0.10434	21
China	316.21	0	0.08228	16

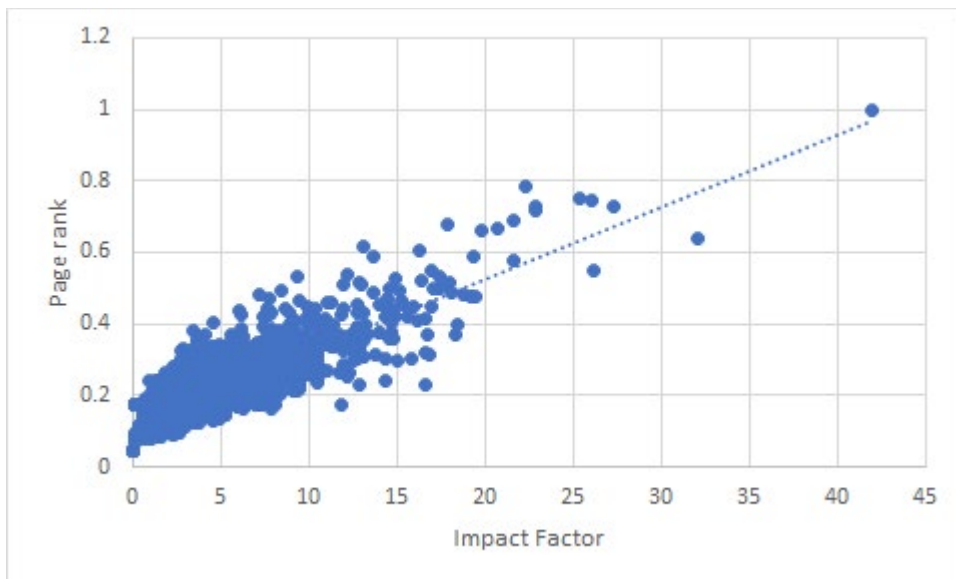


Fig 3 Plot of the Production/Page Rank correlations

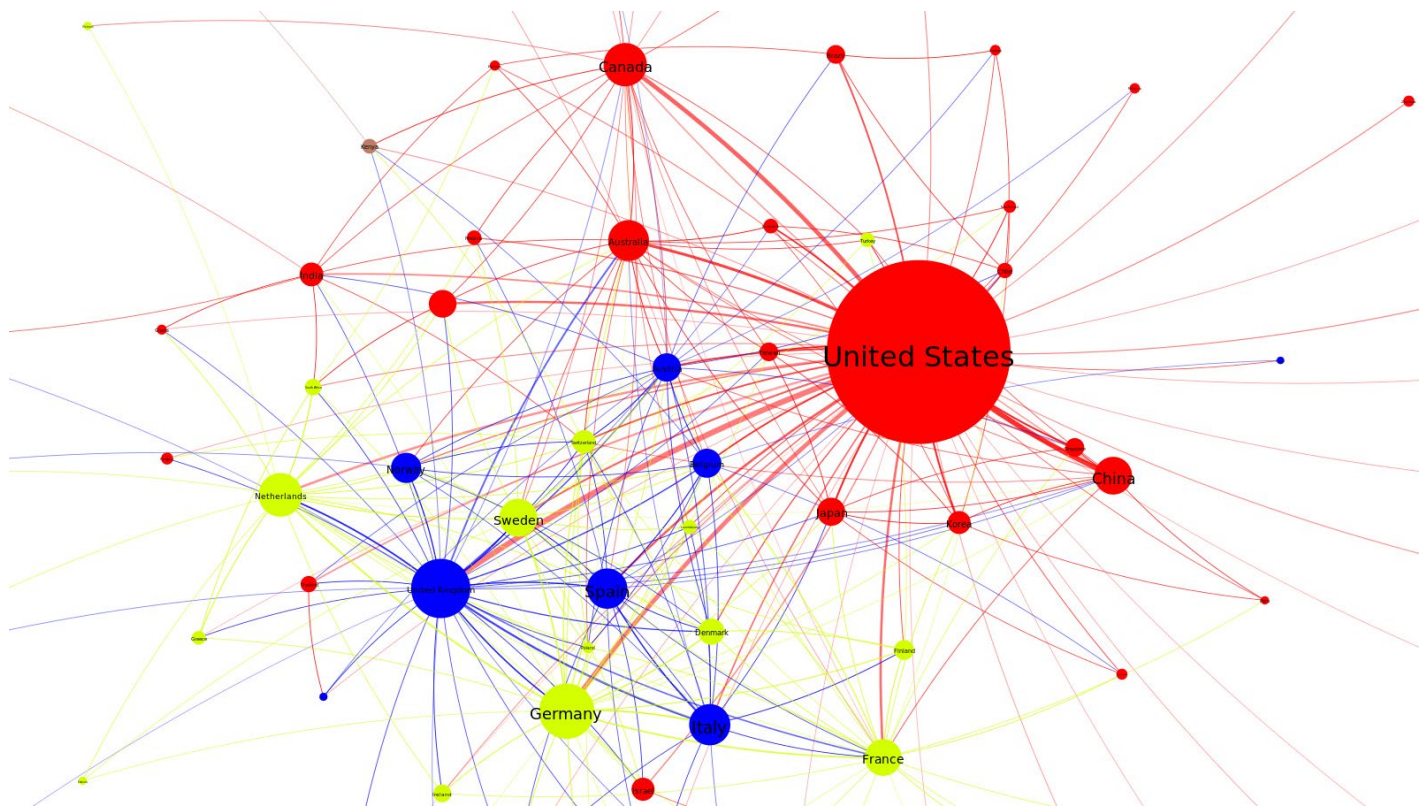


Fig. 4 Country networks

Table 6 Most frequent topics

Total	US	UK	Germany	Canada
India	Children	US	Fertility	US
Fertility	India	Britain	US	Mortality
Children	Mortality	India	West Germany	Age/Cohort
China	Fertility	England	Mortality	Children
Labor supply	Marriage	Wales	Children	Health
Marriage	China	Education	Empirical	India
Mortality	Health	Children	International	Labor Supply
Developing	Developing	Germany	Family	Marriage
Gender	Labor Supply	Cohabitation	Education	Africa

Table 7 Communities

Community	Outcome (IF)	Members
Trude Lappegard (Sociologist)	549.60	152
Hans-Peter Kohler (Economist)	444.155	117
Haoming Liu (Economist)	411.72	99
Samuel H. Preston (Economist)	394.25	81
Joseph P. Price (Economist)	361.98	99
Jeron Van Bavel (Humanities)	333.84	81
S. Philip Morgan (Sociologist)	315.38	74
Catalina Amuedo-Dorantes (Economist)	311.39	98
Kathleen Mullan Harris (Sociologist)	289.71	77
James P. Smith (Economist)		
William G. Axinn (Sociologist)	286.56	99
Greg J. Duncan (Economist)	286.09	62
John Bongaarts (Demographer)	287.33	72
Robert A. Hummer (Sociologist)	274.79	60
Robert F. Schoeni (Economist)	273.54	60
	269.42	74

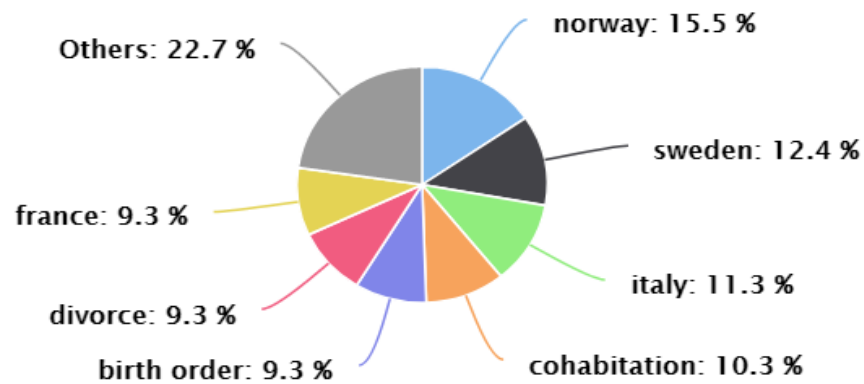
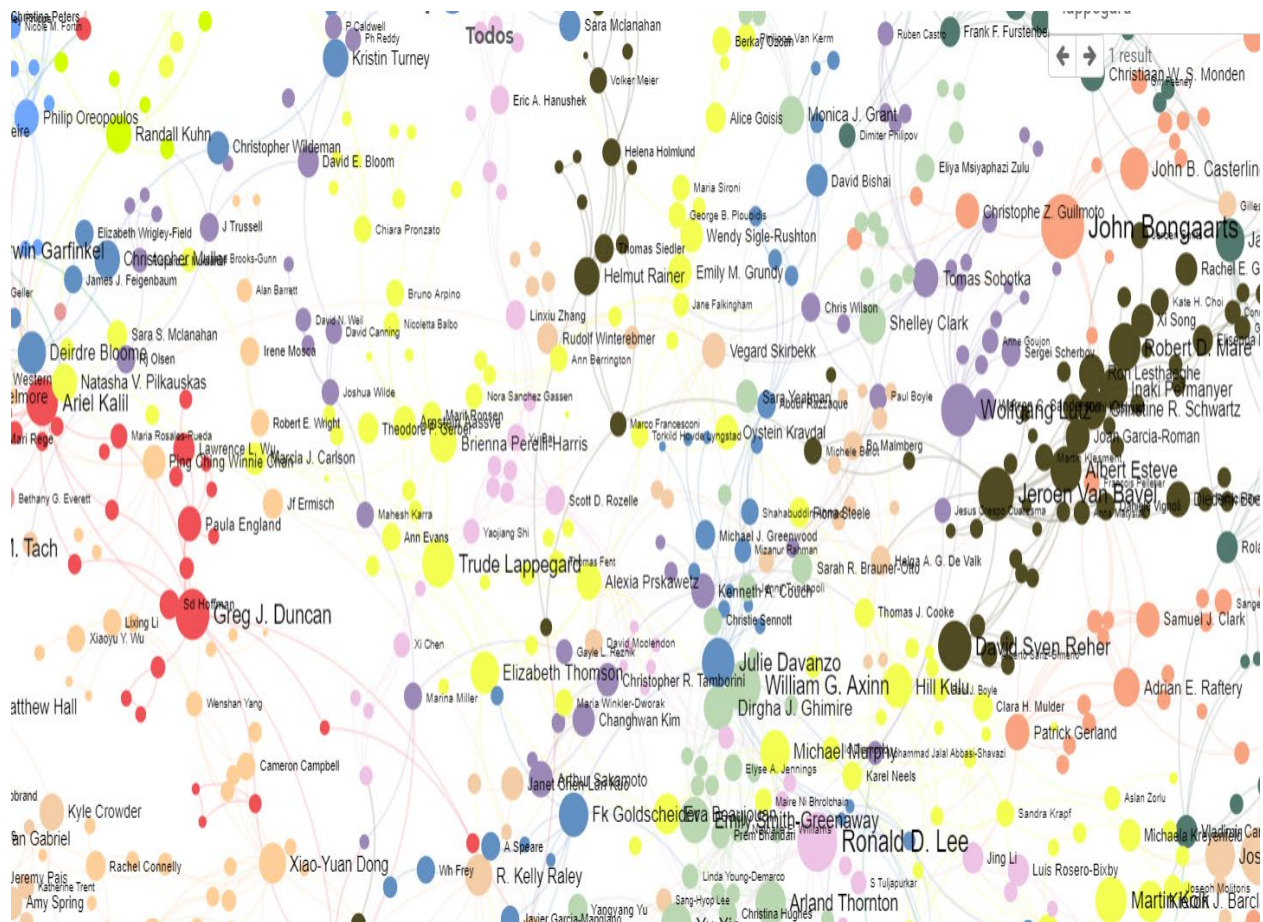


Fig. 5 Lappegard's community

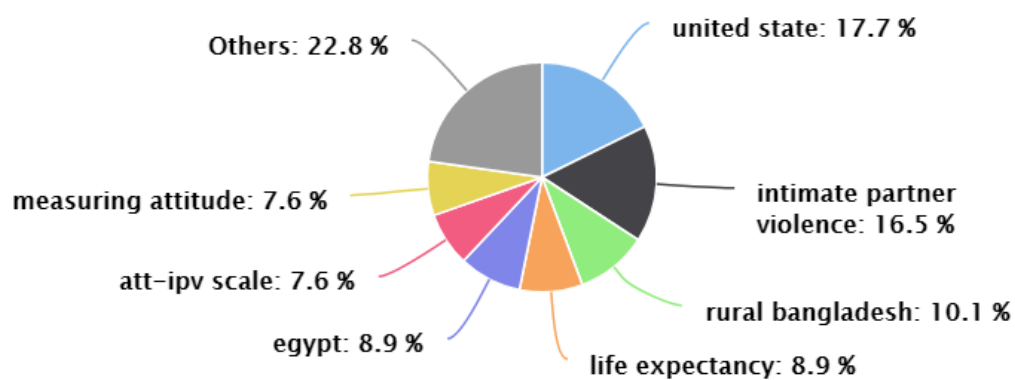
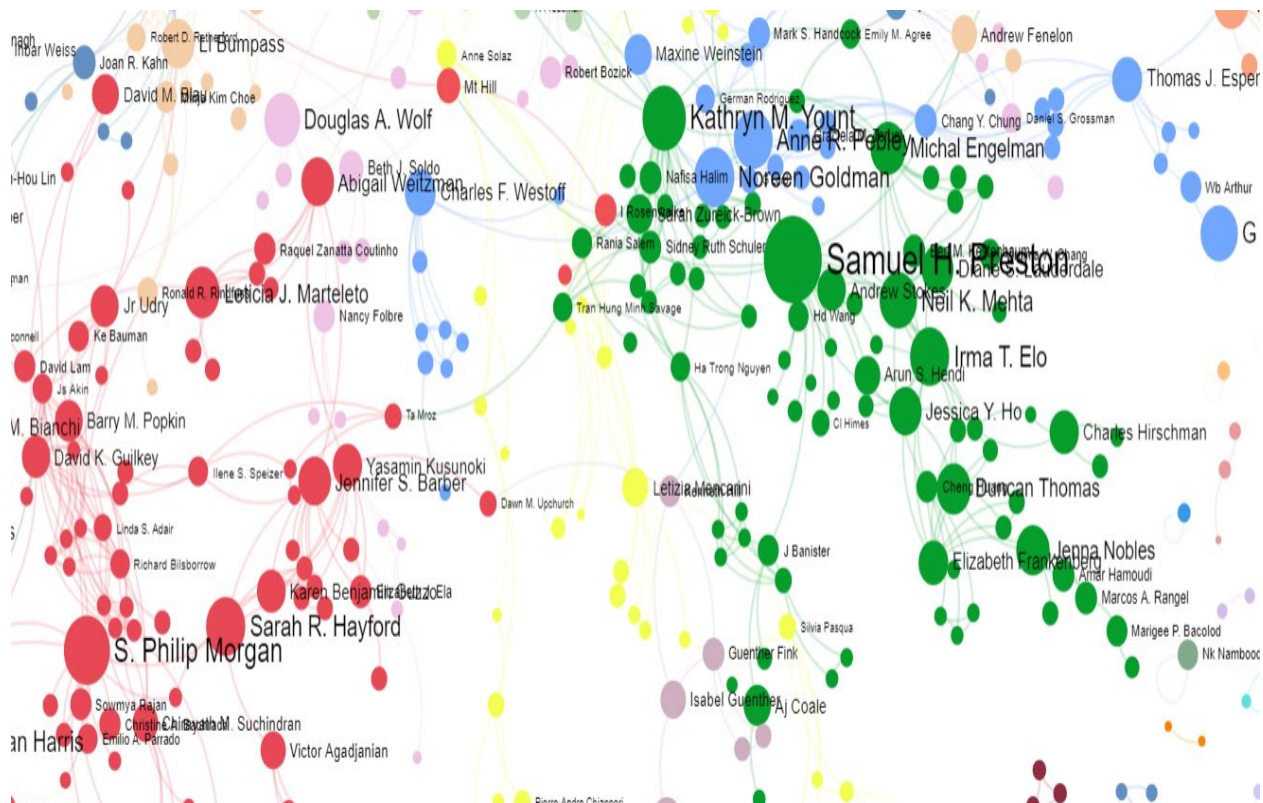


Fig. 6 Network & Keywords of Preston's Community

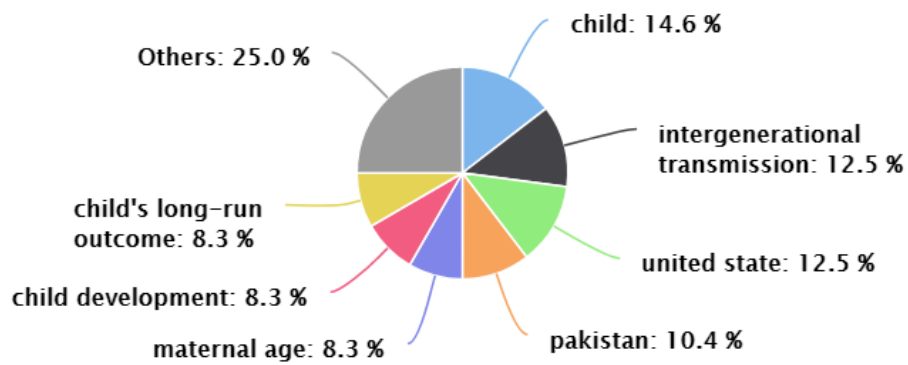
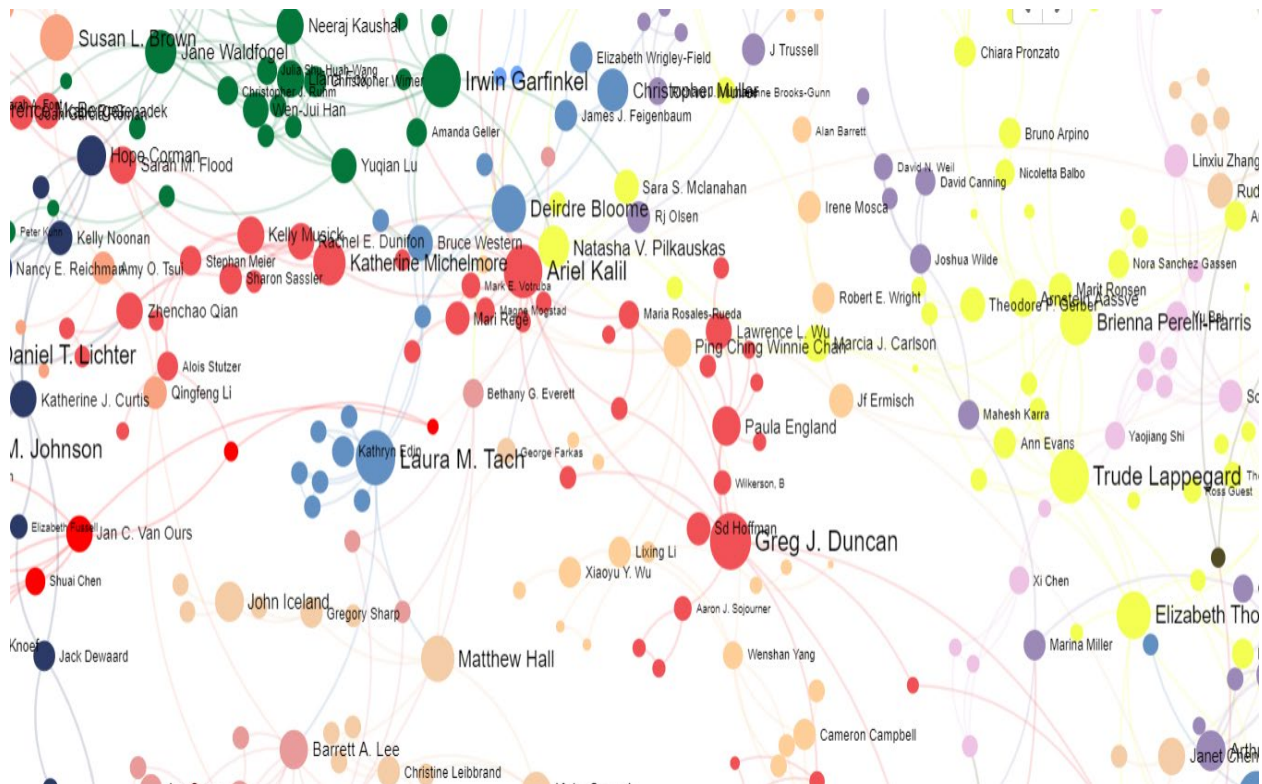


Fig. 7 Network & Keywords of Duncan's Community

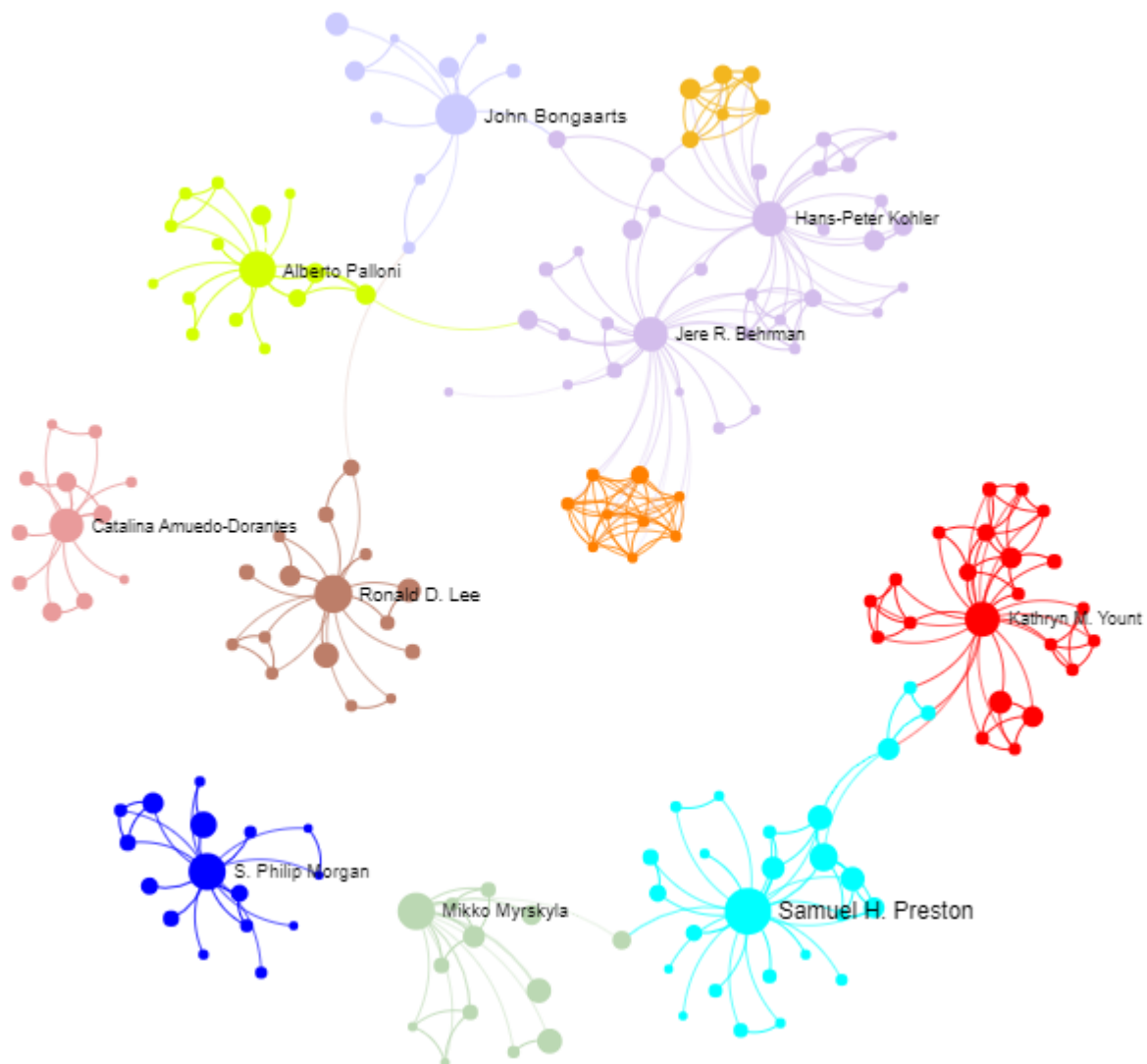


Figure 8 Network formed by the top10 authors according to their IF and all their coauthors, colored by community.

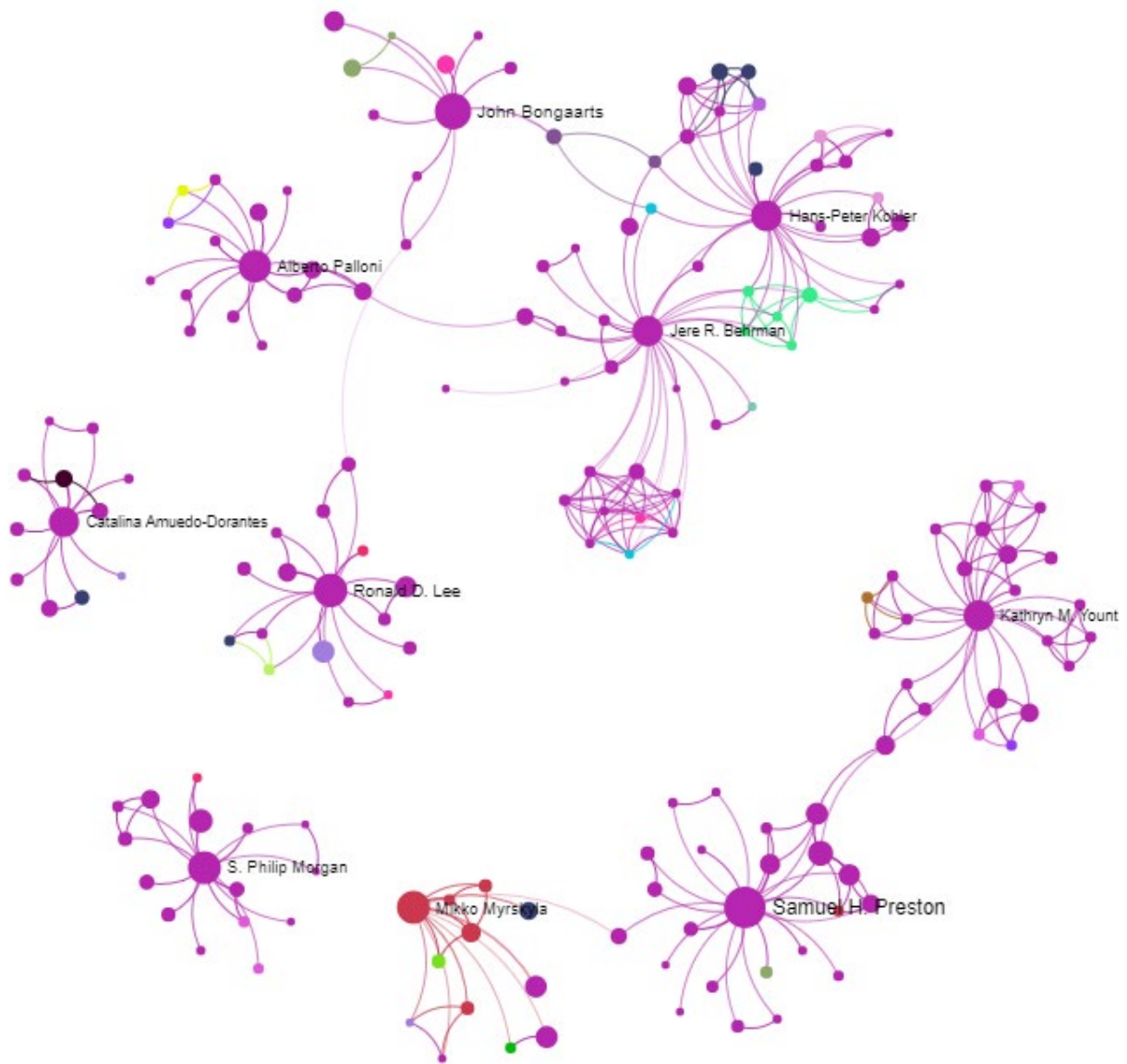


Figure 9 Network formed by the top10 authors according to their IF and all their coauthors, colored by country.

Table 8 JOPE Total IF ranking of authors

	Total outcome	Between-ness	Page Rank	Contributors
Pierre Pestieau (Economist)	11	1	1	12
Gregory Ponthiere (Economist)	7.80	0.25	0.65	5
Akira Yakita (Economist)	7.36	0.04	0.71	3
Klaus Prettner (Economist)	6.41	0.10	0.69	4
Deborah A. Cobb-Clark (Economist)	6.08	0.10	0.66	4
Luciano Fanti (Economist)	5.38	0.04	0.52	3
Zhong Zhao (Economist)	5.34	0.55	0.77	8
Juan Pantano (Economist)	5.25	0.10	0.63	4
Massimiliano Bratti (Economist)	4.98	0.29	0.62	5
Carol Graham (Economist)	4.98	0.14	0.58	4

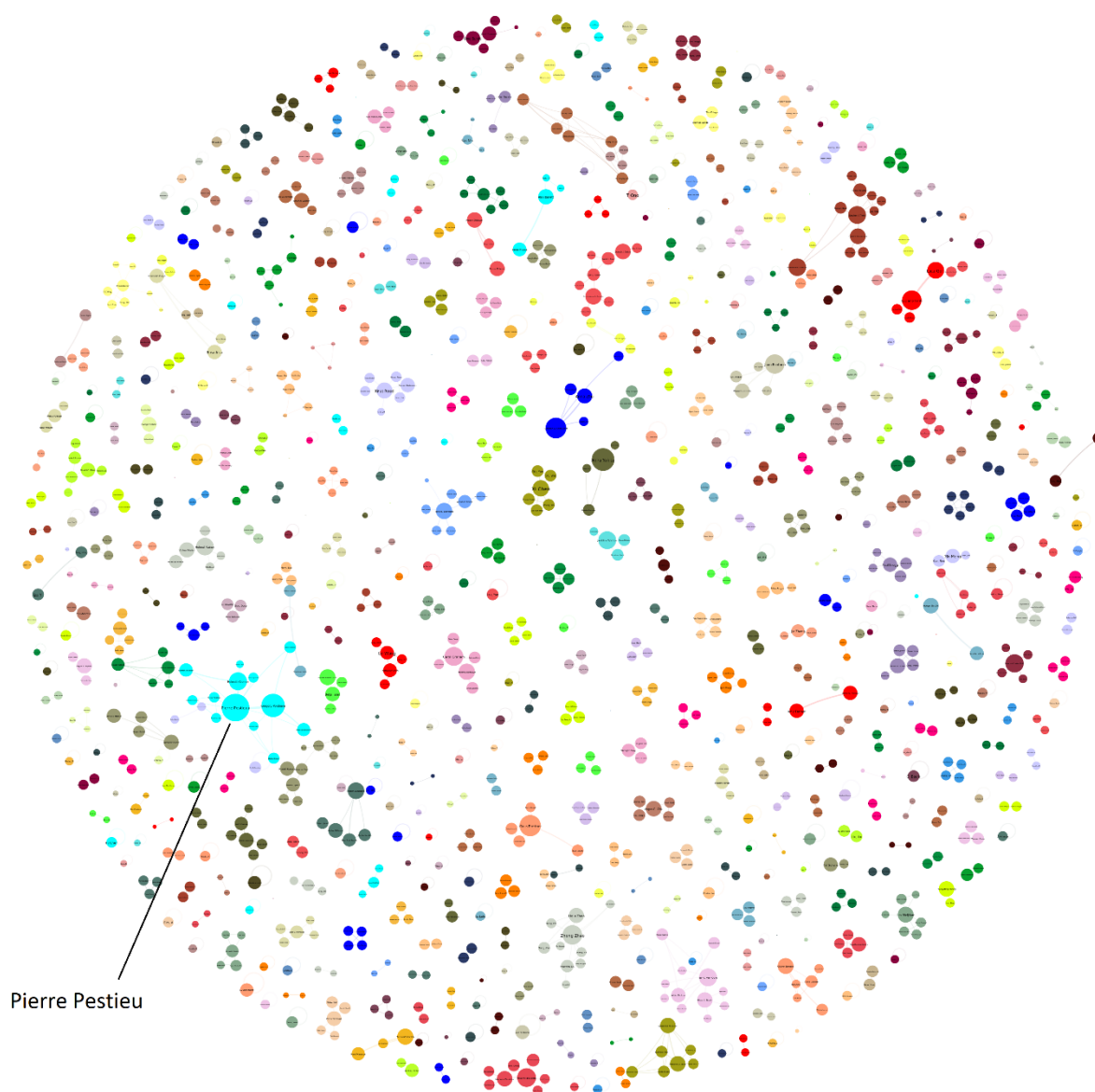


Fig 10 JOPE network

Table 9 JOPE Total IF ranking of authors' countries

	Total outcome	Between-ness	Page Rank	Contributors
United States	380.43	1	1	30
Germany	186.48	0.43	0.46	20
United Kingdom	133.38	0.26	0.54	21
Italy	121.04	0.07	0.27	13
Australia	83.85	0.00	0.24	9
China	74.59	0.06	0.24	6
Sweden	69.43	0.00	0.21	8
France	67.67	0.12	0.40	12
Canada	54.83	0.06	0.27	8
Japan	52.02	0.00	0.19	6

Table 10 JOPE Topics

Total
Fertility
US
Children
Growth
Labor supply
Education
China
Germany
Migration
Intergenerational transmission
Birth order

Table 11 JOPE Communities

Community	Outcome (IF)
Pierre Pestieau (Economist)	38.92
Stephen J. Trejo (Economist)	23.99
Jan C. Van Ours (Economist)	20.96
Zhong Zhao (Economist)	19
Maxximiliano Bratti (Economist)	17.78
Catalina Amuedo-Dorantes (Economist)	17.28
Caron Graham (Economist)	15.47
Lawrence M. Kahn (Economist)	15.08
Mizuki Komura (Economist)	13.60
Deborah A. Cobb-Clark (Economist)	13.47

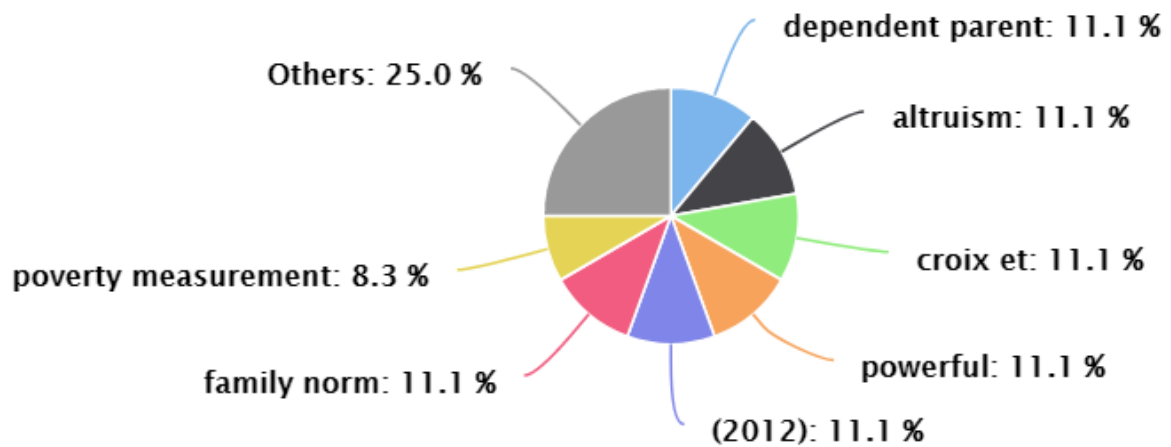
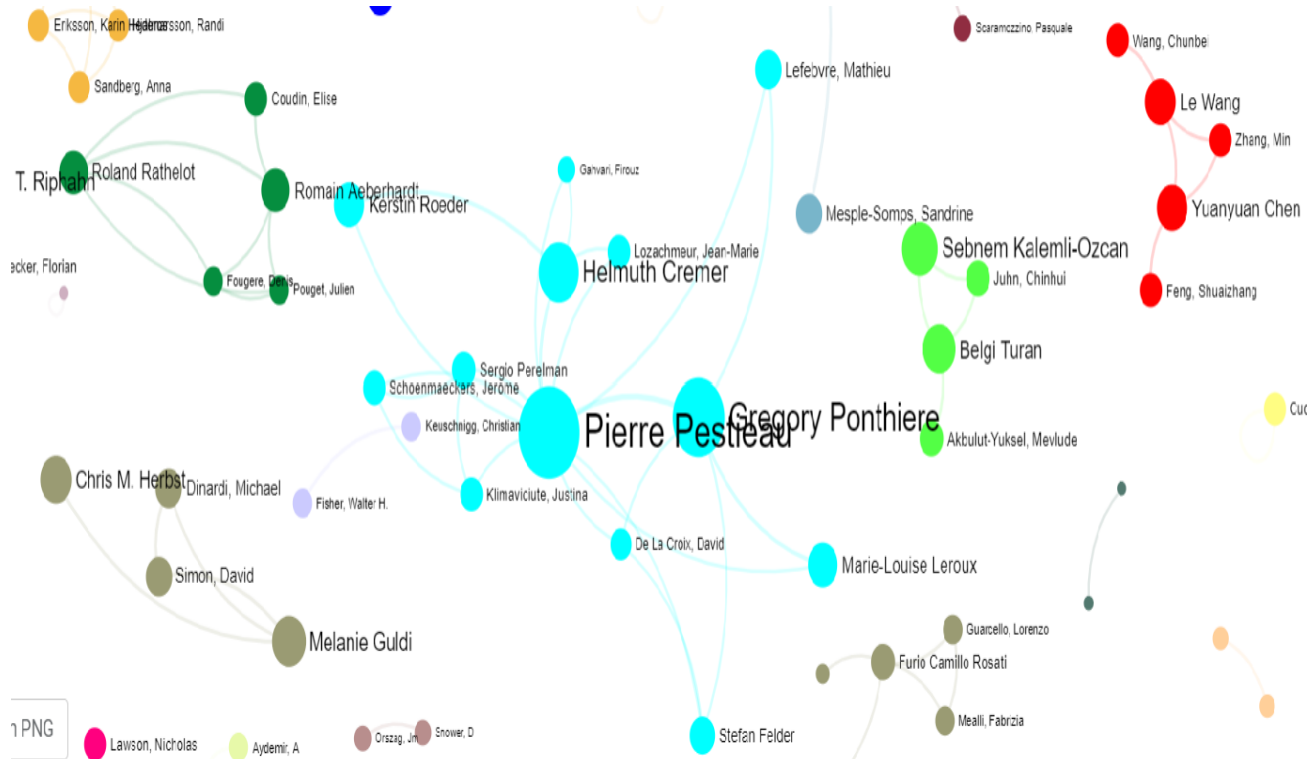


Figure 11 Pierre Pestieu community